MAY 25 1922

Journal of Educational Research

OFFICIAL ORGAN OF THE
EDUCATIONAL RESEARCH ASSOCIATION

B. R. BUCKINGHAM, EDITOR
ASSOCIATE EDITORS
E. J. ASHBAUGH W. W. CHARTERS GEORGE D. STRAYER
GUY M. WHIPPLE S. A. COURTIS LEWIS M. TERMAN

VOL. V MAY, 1922 No. 5

CONTENTS

ACTIVITY ANALYSIS AND CURRICULUM CONSTRUCTION ..... W. W. Charters 357
THE EFFECTS OF SUPERVISED STUDY IN ENGLISH COMPOSITION ..... J. W. Heckert 368
SOME INDIVIDUAL DIFFICULTIES IN THE STUDY OF MUSIC ..... Esther Alku Gaw 381
INTELLIGENCE TESTS AND THE CLASSROOM TEACHER ..... Arthur W. Kallom 389
SCALES FOR MEASURING RESULTS OF PHYSICS TEACHING ..... Harold L. Camp 400
CONVENIENCE AND UNIFORMITY IN REPORTING NORMS FOR SCHOOL TESTS ..... J. Crosby Chapman 406
AN ACCURATE INDEX OF NATIONALITY Riverda Harding Jordan 421
EDITORIALS 426
REVIEWS AND ABSTRACTS 431
NEWS ITEMS AND COMMUNICATIONS 438
EDUCATIONAL RESEARCH ASSOCIATION 450
INDEX FOR VOLUME V, FOLLOWING PAGE 452

Entered as second-class matter at the Post Office at Bloomington, Illinois under the act of March 3, 1879.

Journal of Educational Research

Copyright 1922, by
Public School Publishing Co.

Official Organ of the Educational Research Association

Editor-in-Chief, **B. R. Buckingham**
Director of Bureau of Educational Research, Ohio State University, Columbus, Ohio

Associate Editors

E. J. ASHBAUGH, *Ex-Officio, Secretary of Research Association, Ohio State University.*

W. W. CHARTERS, *Carnegie Inst. Tech., Pittsburgh.*

S. A. COURTIS, *City Schools, Detroit, Mich.*

GEORGE D. STRAYER, *Columbia University, New York City.*

LEWIS M. TERMAN, *Leland Stanford Junior University, Stanford University, California*

GUY M. WHIPPLE, *College of Education, University of Michigan, Ann Arbor, Michigan.*

Contributing Editors

VIRGIL E. DICKSON, *Director, Bureau of Research and Guidance, Public Schools, Oakland and Berkeley, California.*

CHARLES FORDYCE, *Director of Bureau of Educational Measurements, University of Nebraska, Lincoln, Nebr.*

M. E. HAGGERTY, *Dean of the College of Education, University of Minnesota, Minneapolis.*

V. A. C. HENMON, *Director, School of Education, University of Wisconsin, Madison.*

ARTHUR W. KALLOM, *Assistant Director, Department of Educational Investigation and Measurement, Public Schools, Boston, Mass.*

F. J. KELLY, *Dean of University Administration, University of Kansas, Lawrence, Kans.*

ROBT. H. LANE, *Assistant Superintendent of Schools' Los Angeles, Cal.*

WILLIAM A. McCALL, *Assistant Professor of Education, Teachers' College, Columbia University.*

GEORGE W. MELCHER, *Director, Bureau of Research and Efficiency, Public Schools, Kansas City, Missouri.*

EUGENE A. NIFENECKER, *Assistant Director, Bureau of Reference, Research and Statistics, Public Schools, New York City.]*

JOSEPH P. O'HERN, *Assistant Superintendent of Schools and Director of Efficiency Bureau, Public Schools, Rochester, New York.*

LEON O. SMITH, *Assistant Superintendent (in charge of Educational Research) Public Schools, Omaha, Nebraska.*

W. W. THEISEN, *Director, Division of Reference and Research, Public Schools, Cleveland, Ohio.*

CLIFFORD WOODY, *Director Bureau of Tests and Measurements, University of Michigan, Ann Arbor, Mich.*

The JOURNAL OF EDUCATIONAL RESEARCH is published at Bloomington, Illinois, monthly, except July and August by the Public School Publishing Co. The subscription price is $4.00; $4.20 in Canada or Mexico; and $4.50 in other foreign countries.

The editorial office is at Columbus, Ohio. Address all communications on editorial matters to B. R. Buckingham, Director of Bureau of Educational Research, The Ohio State University, Columbus, Ohio.

Subscriptions and all communications about reprints or other business matters should be addressed to the Public School Publishing Company at Bloomington, Illinois.

JOURNAL *of* EDUCATIONAL RESEARCH

Volume V MAY, 1922 Number 5

ACTIVITY ANALYSIS AND CURRICULUM CONSTRUCTION[1]

W. W. CHARTERS
Carnegie Institute of Technology

There are two clear-cut theories concerning the relation of the curriculum to the activities of the individual. The theory upon which current curricula have usually been constructed maintains that the content of the curriculum should come from the specialist and be applied by the individual to the activities in which he engages after he has learned them. The specialist in chemistry, for instance, constructs a course in that subject which presents the fundamental facts and principles of the science as the chemist uses them. The same material is taught to chemists, engineers, farmers, doctors, housekeepers, and laymen, with the expectation that each will apply such of it as he may need, to the activities of his vocation or of his leisure. The lexicographer selects the spelling list for the school and thereby specifies what he considers to be the fundamental spelling vocabulary. This list the individual learns and of it he uses those words which he may need in writing letters, articles, speeches, or reports. To these illustrations may be added the curricula in history, botany, zoology, physics, design, mathematics, grammar, sociology, and others. In all of them the theory of specialistic determination holds that the fundamental content of the subject taught in the school is determined by the specialist; and that the learner, after acquisition, uses what he needs.

[1] A paper read at the meeting of Section Q, American Association for the Advancement of Science, Toronto, December, 1921.

The other theory, which is the natural outgrowth of the theory of evolution, maintains that the content of a school subject is determined by the use to which it will be put. It quite willingly grants that the chemist shall determine the content of chemistry courses for prospective chemists but it emphatically maintains that the same right shall be granted for the use of the engineer, the doctor, the housekeeper, the farmer, and the layman.

The theory of evolution holds that a structure which functions successfully in one situation may be beneficially modified when placed in another situation. The supporters of the theory of activity analysis, to apply the parallel with evolution, hold that the situation of the specialist as described in terms of his activities, duties, and problems, is not identical with that of the engineer, the doctor, or the layman whose problems and duties are different and that, therefore, the structure of the curriculum may well be modified more successfully to cope with the differences of each.

But the method of activity analysis advances beyond the law of natural selection which operates rather blindly in the field of biological life in preserving useful accidental variations. A biological organism changes its structure through accidental variation so that when it comes into possession of characteristics which are of superior efficiency in controlling a type of situation in which it is placed, it tends to retain the useful modifications thus more or less accidentally thrust upon it. Rational beings, however, in developing social organisms, such as the curriculum is, can advance a step beyond the law of natural selection by first making an analysis of the situation to which adjustment is to be made and then developing an organism which meets these conditions in a superior way.

This is primarily what activity analysis attempts to do in the field of curriculum construction. The activities in connection with which a subject may be used are analyzed to discover exactly what the individual is to do and then the subject material necessary to assist in the performance of these activities is collected and organized. For instance, one situation in which the learner

finds himself is concerned with the communicating and recording of facts. Specifically, the laymen writes letters, and children compose original themes. To analyze this situation a count is made of words—in personal letters and in the writing of children. The lists so obtained become the spelling vocabulary for the elementary school, which is the school for laymen. For the spelling curriculum of the business college in which stenographers and typists are trained, the same method is used but applied to a different situation in which word counts of business letters are made. Arithmetic for the layman may be determined by a collection of arithmetical problems which the layman meets. Mathematics for chemists is determined by an investigation of the mathematics used by chemists. The chemistry used in food courses is obtained by derivation from the content of food courses. The grammar necessary for the correction of errors is secured after an analysis of grammatical errors of school children. The English composition needed by engineers is disclosed by an analysis of the uses to which engineers put the subject. The sociology needed by the layman is obtained by an examination of the social problems which confront the layman, just as the sociology needed by the sociologist is determined by the problems encountered in the field of sociology.

In all cases the theory of activity analysis holds that the structure of a subject varies with its function and its content with its use. It leaves a place for the specialists' organization of the curriculum by allowing that the function of the subject for him is one of many functions while it opposes the claim that the structure which is best for the conditions of the specialist meets all other situations equally well.

The method of activity analysis is profoundly significant and is so incompletely understood that the writer has taken the liberty of presenting some of the problems which have to be met in its use. For clarity these will be listed and described as *dichotomous* situations in which the selection of one or another of pairs of considerations modify methods of analysis and of curriculum construction.

Parenthetically, it may be said that an activity analysis is the analysis of both the mental and physical activities which are carried on by individuals. Used in this broad sense, activities include not only what people do but what they think, and feel, or will. In its most complicated form, the analysis is concerned with a broad range of physical actions, feelings, and purposes, as in the analysis of the activities of women made for the purpose of constructing a college curriculum for women. In its simpler forms, it involves the analysis of definite operations, to which the term job analysis is applied, as in the analysis of the operations involved in running a machine. Between these two extremes are activities displaying all gradations of complexity and requiring all degrees of analytical ability.

The first dichotomous situation which causes a modification of methods is contained in the distinction between sequential and non-sequential activities. By a sequential operation is meant one in which the constituent activities normally follow each other in regular order. Most personal routine duties are of this sort, such as dressing in the morning and baking a cake, or as in machinery operations in production plants and the annual calendars of animal husbandry. Non-sequential activities, on the contrary, are arranged in no regular order in the performance of any operation. Specifically, communication of ideas through spelling, writing, and language is non-sequential because there is no regular order observed in the occurrence of words or rules of syntax since their use is dependent upon the demands of varying situations. A job analysis of salesmanship, in which all the activities or duties are to be listed, is much more difficult than the job analysis of cabinet making. A job analysis of home making deals with duties some of which are routine and sequential but most of which are irregular and non-sequential in their appearance.

This leads to very marked differences in methods of analysis. For sequential jobs, the usual procedure followed by the job analyst is to obtain from each of several operators in separate interviews a list of his duties, to combine these, to resubmit the list to them, and to have it checked by the foreman. In non-

sequential jobs the same procedure is used as a basis but from knowledge born of experience, the analyst finds that he needs to make a supplementary analysis by long continued personal observation of the operations and in some cases by a long continued performance of the job. A case in point is that of the analysis of salesmanship which was made by us last year. Interviews were held with many salespeople for the purpose of obtaining a list of their difficulties; to this were added interviews with several managers and superintendents; but so little was obtained from these sources that a member of our staff with competent experience in selling, had to be detailed to watch salespeople at work daily for three months before we felt that we had a reasonably accurate list of the duties of retail selling. In other words, the analysis of non-sequential jobs is much more difficult than that of sequential activities and requires a more elaborate technique.

A second bifurcation of methods is necessary according as the analyst wishes to make an activity analysis or a difficulty analysis. For instance, a spelling vocabulary based upon analysis may list either the words most frequently used in correspondence or those most frequently misspelled. In the former case the top words are such as: *in, all, and, as,* and *for.* While in the latter the words which lead the list are: *separate, which, there,* and *their.*

While theoretically every learned activity at some time or another is a difficulty, with the result that eventually the two lists are identical, there are practical situations in which the activity analysis is not as efficient as the difficulty analysis. For in the first place, operators are often unable to report what they do because many of their duties have become submerged habits, while they may easily be able to recognize the duties with which they have difficulty because difficulties require the focusing of attention. In the second place, if the analysis is made after activities have been partially learned it is extremely important to know those with which the learner now has difficulty in order that special attention may be given to them. So, in the study

of language in the school, to which children come with some proficiency in speech, it is less important to know the rules they use than the rules they break. For practical purposes, in this case, it is unnecessary to teach them rules which they cannot break while obviously it is of first importance to teach them rules whose observance will eliminate error. When the operations have been carried along to a position of partial mastery, difficulty analysis becomes a necessary method in the construction of a training curriculum.

This is the fundamental position of Bobbitt in his theory of curriculum construction, though it can be demonstrated in opposition to this position that it will hold only where activities have been partially learned and is, therefore, not of sufficient breadth to determine all the activities which should be included in a school curriculum. These can be determined only by an activity analysis, to which a difficulty analysis is a necessary supplement in the later stages of learning.

A third distinction of sufficient importance to be mentioned is that between what may be called gross analysis and detailed analysis. If the analysis is made sufficiently minute it is possible so objectively to determine the content of the curriculum that the work can be performed in the laboratory; but if the analysis ceases before working units are obtained, it becomes necessary to supplement the findings of the analyst by conference with faculty committees. For instance, a study of the mathematics used in chemistry by Williams is carried to sufficient detail for not only the operations but their range of complexity to be determined with such exactness that a review course may be worked out in detail. At no point in the study is any item a matter of opinion. So, also, the spelling lists obtained by word counts are sufficiently detailed to establish objectively the actual content of the spelling curriculum. On the other hand, the job analysis made by Strong of the duties of executives, as the basis for a curriculum in commercial engineering, was so complicated and the pressure of time so great that only the broad outlines of the content of the derived courses could be determined, with the result that for the comple-

tion of the work committees of teachers of all subjects concerned become a necessary part of the method.

Gross analysis, then, is not so objective as detailed analysis, and requires the use of concensus and conference while a detailed analysis makes opinion an unnecessary part of the methods of curriculum construction. The goal of all curriculum construction is the elimination of as much of the subjective element as possible, though this cannot be entirely accomplished.

A fourth distinction in the use of activity analysis arises from divergent uses of the analysis. The employment departments of commercial and industrial organizations make analyses to obtain job specifications while the training departments use them for curriculum content. For job specification, it is necessary to list only what is to be done by an operator in a specific position so that when the employment office is called upon to select a person for the position he will know the nature of the job. For curriculum construction, it is necessary to supplement this information about what is to be done with additional information as to how the operations are to be performed. For these reasons the job specification is much less detailed, and as a matter of fact it is usually found that the curriculum cannot be based upon the employment department's job specifications. New analyses have to be made by the training department with much greater attention to the details of the operations. Theoretically, the job specification is the proper basis for the training program; but because of these differences in function those job specifications found in employment offices will not usually suffice.

This leads naturally to a fifth distinction which may be stated in the form of a question. Shall training departments teach the installed methods of job performance or the best methods? The obvious alternative is the latter, but practical considerations may cause a change of opinion. For instance, if in the training of operatives in the correspondence division of a large department store the management asks the training department to make a change in the system as well, it is in effect asking it not only to teach but to reorganize. So, because of the expense and time

involved, it may be advisable for the training department to teach the imperfect installed procedure rather than to effect a more efficient procedure before the training program is begun. Yet, it may be said parenthetically, that in business the tendency is toward placing both training and organization on the shoulders of the training department, especially when a radical reorganization is not required.

Nor are the schools free from this problem. Superficially, it seems self-evident that only the best methods of performing activities should be taught in the school. But in practice it is often a question whether this is wise. Is it advisable, for example, to advocate the best methods of solving social and political problems (the best being determined by the judgment of conscientious teachers) if it is not practically possible to make the necessary reorganization of social institutions which will permit the most efficient methods to be used? In these difficult cases the wise teacher modifies his best in such a way as to make it the best possible under existing conditions and teaches his students the methods by which reorganization of social institutions may be effected. But in either case the curriculum is affected by the point of view which is chosen.

The sixth consideration is the distinction between ideals and activities. An activity analysis is not sufficient to determine the curriculum unless standards of performance are set up by which to make a selection of methods. This is done by the determination of ideals, which dominate the performance of the activities. This is recognized informally in the selection of men in business by the inclusion in the job specification of both the duties to be performed and the personal qualities desired. Some positions require courtesy, others demand scholarliness, and all require honesty and conscientiousness. So in building a curriculum it is necessary to determine both activities and ideals. For instance, a secretary may have the duties of answering the telephone and of meeting callers. These may be done in several ways but obviously the ideal of courtesy is so important that only courteous methods of answering the telephone and of meeting callers

will be selected. Merely to perform these duties, or to perform them with wrong ideals, will not lead to efficiency.

Consequently in school as in business the determination of standards in the guise of ideals is an essential part of all curriculum construction. So important is this procedure that faculties should conscientiously and reflectively determine the ideals whose stamp their graduates shall bear and then apply the appropriate ideals as standards for the selection of alternative methods of performance. Long and laborious as this procedure is, it is the only method by which the curriculum can be built or can be acquired. In all curriculum construction, the ideals must be determined before the methods for performing the activities obtained by analysis can be selected.

Of a somewhat different sort is the seventh bifurcation. This lies in the distinction between primary subjects and derived subjects. To clarify this distinction an illustration may be given. When in the major job analysis made by Strong the duties of the executives of certain industrial organizations had been ascertained and the information necessary to perform them had been listed, there was found to be a large amount of English, of psychology, and of economics, but only a very short list of the principles of physics. So far as the primary subjects, that is those discovered in the analysis, were concerned physics was judged to be relatively unimportant; but this did not mean that no other items of physics were necessary. For, when the primary subjects, such as mechanics, were examined, it was found that physical laws were necessary for their mastery. So, to obtain a full curriculum in physics it was necessary to list all the items of physics revealed in the primary analysis and to derive from the other subjects the additional items of physics necessary to understand them.

This means that the so called fundamental service subjects of mathematics, chemistry, physics, sociology, or history are essentially derived subjects rather than primary subjects revealed in the analysis. It may be said in passing that this explains in part the statement of professional men that they do not use their

fundamental subjects and explodes the claim sometimes made that because they do not use them, they should not learn them. For, obviously, at one time it was necessary for them to learn the fundamental subjects in order to understand methods which they now use; and it is legitimate to forget items once the basic use for them has ceased. However, this does not mean that the method of specialistic determination is to be followed. It is still necessary for the fundamental service subjects to be derived by activity analysis so as to include only what must be understood; but it is not a matter for alarm if later they are forgotten after use for them has disappeared.

Finally, a distinction needs to be made between the vocational elements and the extra-vocational elements in the curriculum of the vocational school. The vocational portions of the curriculum can be obtained by the analysis of vocational activities and the material may be collected under the domination of selected ideals. But just as the doctor or the engineer is something more than a mere doctor or engineer because he has extra-vocational responsibilities and interest in his family, his friends, and his civic problems, so the curriculum of the vocational school must include more subject material than is obtained from a job analysis of the vocation.

Two analyses are necessary, a vocational analysis and an extra-vocational analysis. The activities and interests of extra-vocational life, those problems and interests which all men have, irrespective of vocation, must be analyzed by the same methods as are used in the vocational analysis.

At present the extra-vocational content is determined negatively. A few hours are set aside in the course of study for so called cultural subjects and the selection of courses to fill them is largely left to chance. What we advocate is a positive approach to the question through an activity analysis of extra-vocational life from which can be definitely determined what will be of most use for the individual in furthering those interests and needs which are not directly connected with the vocation.

Other considerations might be presented to illustrate the problems of curriculum construction but enough have been mentioned to show the possibilities and the difficulties of activity analysis.

In conclusion, the foregoing considerations demand that when activities are analyzed, ideals and standards must be set up to aid in the selection of subject material; that in analyzing activities, difficulty analyses may supplement activity analyses; and that in all cases the analysis should be as detailed as possible in order to secure the largest measure of scientific objectivity. Nonsequential jobs are difficult of analysis but the task is neither impossible nor impractical. The content of subjects may be of direct use in activities or may be derived from subjects which are of direct use, as the basis upon which they rest. Finally all courses of study in vocational schools must include not only vocational but also extra-vocational content, both obtained by the analysis of activities—the former by an analysis of the vocation and the latter by an analysis of extra-vocational life.

THE EFFECTS OF SUPERVISED STUDY IN ENGLISH COMPOSITION

J. W. Heckert
Miami University

Teaching children how to study has come to be considered one of the important ends in the educative process. It has taken its place by the side of useful information, skills, ideals, attitudes, and prejudices, as deserving of the efforts of the teacher. We have learned in our elementary and high schools, as in our best graduate and technical schools, that the test applied to the learner should not be merely, "What does he know?" but rather, "What can he do independently of the teacher?"

Various plans have been devised for the purpose of teaching children to study. The present investigation concerns itself with the efficiency of the so called divided or double-period plan, by which the first half of a period is given to class discussion and assignment while the second half is given to supervising the work of children upon the new assignment. The modification used in this experiment consisted in the fact that the teacher did not definitely adhere to so rigid a division of time for recitation and study purposes respectively, but called children to a class discussion whenever there was something worth while to discuss, and permitted them to work or study the remainder of the time.

Many claims have been made for the efficacy of the double period plan. Some of these are probably well founded, others are not, for in some instances the judgments are based upon careful investigations; in other cases the opinion of teacher or superintendent is regarded as sufficient proof of the efficacy of the plan. One of the more recent investigations of the problem of supervised study in the divided- or double-period plan was made by Professor Breed of the University of Chicago and others, under the auspices of The Michigan School Masters' Club. Professor Breed investigated the effect of supervised study "when carried

on by this plan under actual school conditions in thirteen high schools of Michigan and one of Minnesota with the aim of throwing additional light on the effectiveness of a common form of organization of supervised study." Breed's conclusions are stated by him as follows:

"1. On the basis of average results for whole classes, supervised study of the type tested was slightly less efficient in first-year algebra, and much less efficient in ninth-grade composition, and was more efficient in first-year Latin, than non-supervision.

"2. Supervised study of the kind tested facilitated the progress of the poorer pupils, and to a correspondingly greater degree retarded the progress of the better pupils.

"3. The divided- and double-period plans should not be urged for general adoption in secondary schools until their efficiency is more clearly demonstrated.

"4. Until the divided- and double-period plans meet the objections urged against them, a selective or differential plan of study-supervision should be favored by secondary schools, that is, a plan which concentrates upon pupils of less ability.

"5. The technique of supervising the study of the poorer pupils needs to be improved; a technique of supervising the study of brighter pupils needs to be developed."

It seemed to the writer, however, that Breed's method of procedure was open to objection in some particulars. Classes were divided upon the basis of previous attainments in English composition as indicated by teacher's grades (semester) and a preliminary test. Under such a plan it was impossible to make allowance for the very capable child who writes a rather poor composition and the slow, plodding child who, because of very constant application has succeeded in attaining a comparatively high score but whose further progress is limited by his lack of capacity. Again the basis of Breed's statistics are teachers' marks unaided by any scales. It would seem that the rating of the papers could have been materially improved by means of one of the several composition meters now available. In the third place it would seem that little was done in this investigation

to apprise the teachers participating of the real significance of supervised study. To be sure the teachers were referred to Whipple's Manual, but this very reference implied a misconception of the problem involved in teaching children how to work more intelligently on their compositions, as will be shown later, and the teachers could receive but little assistance upon a point on which it was most needed to make the investigation a success. Breed agrees with Parker when he holds that the poor students appear to profit most from supervised study and the bright pupils are actually retarded by this method. However, it seemed to the writer that a capable teacher who studies the special needs of the bright child must be able to teach him better methods of work; if she cannot, why should such children be required to attend class exercises at all the very purpose of which always is to teach better methods of work? It would seem that the loss of the bright child in the supervised study period is to be accounted for rather by the inability of the teacher to understand his needs and therefore to deal with him intelligently.

The object of the present investigation was to ascertain (1) whether supervised study in English composition under the divided-period plan or its modifications, when the teacher knows the technique of directing study in this subject, that is, when she knows what to do with the children of varying ability in order to improve their methods of writing compositions, whether such a plan is less efficient than unsupervised study of this subject; (2) whether the poorer pupils necessarily profit most from supervised study and the brighter pupils not only profit less but are actually suffering from the interference resulting from supervised study; (3) what are the best methods of teaching children to improve their methods of work in writing English compositions?

This experiment was conducted with the ninth grade of the William McGuffey High School at Miami University. The National Intelligence Test had been given at the beginning of the year. The coefficient of brightness was computed for each child. The entire class was also given two preliminary tests in compo-

sition writing to ascertain the initial attainments of each pupil. The papers were rated by means of the Thorndike-Hilligas scale. The class was then divided into two groups each of seventeen pupils of approximately equal ability, constituted as follows:

The supervised group had coefficients of brightness totaling 1766 points and composition scores totalling 938.4 points; the unsupervised group had coefficients of brightness totalling 1707 points and composition scores totalling 990.8 points. The advantage in intelligence, therefore seemed to be slightly in favor of the supervised group while the unsupervised group had the advantage in initial attainments in composition writing.

The two groups were taught by the same teacher, Miss Marguerite Royal, staff teacher of English. Each group worked for twenty-five periods of one hour each. Of these periods seven were given during December and January 1920-1921, one period a week; eighteen were given during April and May, 1921, the two groups alternating from day to day. Illness of the teacher interrupted the experiment during February and March. The writer assisted the class teacher in directing the work of the supervised group, primarily for the purpose of developing a method of procedure; the teacher, however, had charge of the recitation in the supervised as in the unsupervised class. In both groups, the teacher attempted to arouse as much enthusiasm for the work as possible in order that every child might do his best. In this she succeeded for we never had keener interest in the writing of compositions.

Our principal task was to devise a method of procedure for directing the work. We decided first of all, to attack the more fundamental aspects of composition writing, how to choose the subject for a composition, how to plan the materials to be used in the development of the subject, how to arrange these materials, what an outline is for, and how to word it and use it. The chief emphasis throughout was placed upon having something to say and saying it in such a way that it would be clear and interesting to an audience. Occasionally a half hour was set aside for the purpose of enjoying the stories written. In order to teach chil-

dren to be more successful in criticising their own productions they were encouraged, first, to go over their compositions in the light of the principles they had been taught, and second, to meet with one or two of their classmates for the purpose of going over their compositions.

The various phases of composition writing were then further analysed, with a view of showing the best methods of procedure of the skilled writers. The effects of this analysis were (1) to make the teacher conscious of the factors in good methods of writing, and (2), as a result of (1), to devote herself more to the process of children's writing rather than the product.

Let me give an illustration to show how we attempted to meet the needs of the individual children in the supervised group. One of the boys chose as his subject "Mischievous Billy." Although he had been taught how to make an outline, there was nothing in his outline to indicate what he planned to say. When asked he said, "I am going to tell how Billy went up town and a lady asked him, 'How do you do Billy,' and Billy said, 'I do as I please.' When the writer objected that this would not do, the boy said that he did not know what else to write. When asked he replied that he had had no one in mind when he chose his subject.

Further questioning brought out the fact that there was a little boy in his neighborhood, who might well be called "Mischievous Billy." This little fellow had visited the yard of the pupil where he had found a coaster wagon and a can of paint; he had proceeded to paint the wagon and had gotten considerable paint on himself. The boy was asked whether he knew another escapade of this little fellow. It happened that this small boy had had some difficulty with children in an adjoining block. The older children took up the fight and punished him by throwing a basin of water on him. The pupil agreed that this would make another good incident for a story about Mischievous Billy. He was then led to recall an incident in which he and his brother were the actors, an incident that might well be added to this story. These two boys had a bantam rooster, a rather troublesome animal

to members of the family and the neighborhood. One morning they tied a tin can to the rooster's tail. The immediate results convulsed with laughter both the boys and other members of the family. The boy thought this incident would do very well for his story. He was then asked to make out an outline. This is his outline: Billy has some fun with a rooster. 2. Billy paints a wagon but also himself. 3. Billy receives a bath. 4. Billy has a settlement with his mother. Next came the writing of the story. The boy showed that further development of the incidents was necessary, because he was too much inclined to state mere facts and did not appreciate the value of interesting details. He was then requested to imagine the various things that might have happened in the incident with the rooster (he had rather completely forgotten them); these he was asked to state orally first. Occasionally the writer offered suggestions, some were received gladly, others were rejected; in no case were they forced upon the boy. When the incidents had been thought through in this detailed fashion, the boy seemed ready to write. Another difficulty now manifested itself. Since he had never taken any interest in composition work he had a rather meager vocabulary. Suitable modes of expressions had to be sought; but after consideration of a number of expressions he had used, he began to attend to this matter upon his own responsibility. Here is the first paragraph from this boy's final test composition. "It was one dismally sultry, rainy day last summer in Wisconsin. It has been raining hard all the night before and of course everything was covered with water. The roads were bad all around the hill but none as bad as the newly made dirt road coming down the hill." This paragraph still is faulty in some respects but it does show a very definite attempt to choose words which are adequate to the ideas to be conveyed, and is immeasurably better than this boy's work six weeks previous. Limitations of space prevent the giving of other illustrations.

Two tests were given to ascertain the improvement which the children had made during the period of the experiment. A list of varied subjects was given the children to choose from and an

hour was given them in which to prepare and write their papers. The pupils were permitted to choose their subjects in advance of the day on which the final test was written. This was done in order to give them a better opportunity to plan their compositions but there was little evidence that the children availed themselves of this opportunity. When the papers had been written, they were numbered according to a scheme which concealed the identity of the writers. The papers of the supervised group were collected with those of the unsupervised group and rated with the Thorndike-Hillegas scale by the staff teacher and the writer. The scores assigned the papers by the two readers were in rather close agreement. Thus in the final set, there was complete agreement in twenty-three out of thirty-three cases, one of the children of the supervised group having dropped out during the spring. In the remaining cases, the differences ranged from three to five points. For the purpose of computing the progress of children averages of the scores of the two readers were used only, and as the final scores, the average scores for the two final tests were used.

The Results of the Experiment

Do children profit from supervised study in English composition, when the teachers know what to do with children during the supervised study period in order to improve their methods of work? Table I furnishes an answer to this question. According to this table, the median gain for the supervised group is eight percent upon the initial ability as indicated by the scores from the preliminary tests, with a median deviation of 6.5. The median gain of the unsupervised group is three percent and the median deviation 6.62. This result is strikingly different from that obtained by Breed, who registers a loss for the supervised groups in English composition.

When we come to our second question, do children of superior ability profit less from supervised study or actually lose as a result of it, we find the answer in Table II.

TABLE I.—PERCENT GAIN OF EACH PUPIL OF THE SUPERVISED AND THE UNSUPERVISED GROUP FROM THE PRELIMINARY TESTS TO THE FINAL TESTS. SCORED BY THE THORNDIKE-HILLEGAS SCALE

Supervised Group			Unsupervised Group		
Number of Pupil	Coefficient of Brightness	Gain percent	Number of Pupil	Coefficient of Brightness	Gain percent
S17	122	21.8	16	99	19.4
S22	114	19.9	29	122	12.8
S-3	109	17.1	6	88	8.9
S-5	89	14.8	15	101	8.0
S19	104	10.9	25	96	7.8
S20	140	10.3	3	107	7.4
S 7	113	9.9	30	83	6.0
S26	98	9.1	4	97	3.2
S13	100	7.0	2	93	3.0
S28	98	6.2	12	103	2.1
S10	88	4.5	9	115	0.0
S31	92	2.6	32	110	0.0
S33	102	-3.0	34	113	-0.7
S14	97	-4.0	1	86	-6.8
S21	91	-6.2	18	99	-7.6
S 8	110	-8.5	11	83	-9.1
			24	118	-16.3
Median Gain.....		8.0	Median Gain.....		3.0
Average Gain.....		7.0	Average Gain.....		2.6
Median Deviation.....		6.5	Median Deviation....		6.6

In Table II the pupils ranking above and below the median in capacity as indicated by coefficient of brightness in each group are compared with reference to improvement made. Here we find that the bright members of the supervised group attained an improvement of 4.8 percent as compared with 1.66 percent for the corresponding numbers of the unsupervised group. The slower members of the supervised group achieved an improvement of 4.27 percent while the corresponding unsupervised group improved 2.72 percent.

If we take initial attainment as the basis of comparison, we find that the fifty percent of the supervised group making the best average scores in the tests improved 3.65 percent, while the corresponding fifty percent of the unsupervised group improved but .16 percent. The remaining half of the supervised group gained 10.37 percent, while the corresponding half of the unsupervised group gained 4.6 percent.

TABLE II.—SHOWING GAIN PERCENT OF UPPER AND LOWER 50 PERCENT BASED UPON COEFFICIENT OF BRIGHTNESS

	Supervised Group			Unsupervised Group		
	Number of Pupil	Coefficient of Brightness	Gain percent	Number of Pupil	Coefficient of Brightness	Gain percent
Upper Fifty Percent	S17	122	21.8	29	122	12.8
	S22	114	19.9	15	101	8.0
	S23	109	17.1	3	107	7.4
	S19	104	10.9	12	103	2.1
	S20	140	10.3	9	115	.0
	S 7	113	9.9	32	110	0
	S33	102	-3.0	34	113	- .7
	S 8	110	-8.5	24	118	-16.3
	Average gain		4.8	Average gain		1.66
Lower Fifty Percent	S 5	89	14.8	16	99	19.4
	S26	98	9.1	6	88	8.9
	S13	100	7.0	25	96	7.8
	S28	98	6.2	30	83	6.0
	S10	88	4.5	4	97	3.2
	S31	92	2.6	1	86	-6.8
	S14	97	-4.0	18	99	-7.6
	S21	91	-6.2	11	83	-9.1
	Average gain		4.27	Average gain		2.72

TABLE III.—COMPARING RATE OF PROGRESS OF CHILDREN ABOVE MEDIAN IN SCORES OF PRELIMINARY TESTS WITH THOSE BELOW MEDIAN

	Supervised Group				Unsupervised Group			
	Number of Pupil	Average Score	Coefficient of Brightness	Gain Percent	Number of Pupil	Average Score	Coefficient of Brightness	Gain Percent
Upper Fifty Percent	S31	68	92	2.6	24	72	118	-16.3
	S20	65	140	10.3	29	66	122	12.8
	S 8	65	110	- 8.5	18	65	99	- 7.6
	S13	61	100	7.0	4	63	97	3.2
	S28	61	98	6.2	34	62	113	- .7
	S14	61	97	- 4.0	25	61	96	7.8
	S21	60	91	- 6.2	12	61	103	2.1
	S17	59	122	21.8	9	60	115	0
	Average Gain			3.65	Average Gain			0.16
Lower Fifty Percent	S33	58	102	- 3.0	32	60	110	0
	S10	55	88	4.5	2	58	93	3
	S26	55	98	9.1	6	56	88	8.9
	S19	55	104	10.9	11	55	83	- 9.1
	S22	54	114	19.9	15	53	101	.8
	S23	51	109	17.1	16	51	99	19.4
	S 7	51	113	9.9	3	51	107	7.4
	S 5	49	89	14.8	1	51	86	- 6.8
					30	46	83	6
	Average Gain			10.37	Average Gain			4.6

Table III comes nearest to supporting the inferences of previous investigators of this problem: i.e., children with high initial attainments under supervised study organization show but little improvement while those of low initial attainments show relatively high gains. But even these tables do not show actual losses for the supervised bright child and high gains for the unsupervised bright child. Is it possible that children in former investigations of the efficiency of supervised study were using a larger amount of time for their work than those of the supervised groups? Is it further possible that they received assistance in their homes which really made the work of the unsupervised groups supervised work? This experiment was controlled in order to prevent the assistance which children might receive in their homes and their use of more time for their tasks than the supervised group.

These tables also point out another possible difficulty with previous experiments. The gain or loss in these tables is based upon the score which children made in the preliminary tests, upon ability in the subject in which rate of progress was to be measured. Is it possible that the actual attainments in a given subject have been mistaken for capacity for achievement in that subject? By actual attainments, as indicated by the preliminary tests, seven out of sixteen children were misplaced, that is, they were classed with the pupils of low capacity when they should have been ranked with the superior. The pupils in question are Numbers 33, 19, 22, 23, 7, 15 and 3. It is of course possible that English composition is peculiar in this respect, that capacity correlates less closely with actual attainment than in other school subjects. Thus in the above group there were at least four boys who, while very capable in many lines, were not interested in writing compositions. It may also be objected that an intelligence test is not an absolutely sure indication of the intelligence of the pupils tested. This objection, however, would not hold in the present case, for in passing upon the validity of the division made there were available the results of other tests previously given and the teachers' estimates of the children's ability. The division into two groups,

therefore, was probably as accurately made as possible. It would seem, therefore, that this is a point which requires attention in future investigations involving the division of classes into groups of equal capacity.

Summary

Our results seem to indicate that carefully directed study or work of children in writing compositions is decidedly valuable to children receiving such supervision.

The results of this investigation further indicate that under fairly skillful direction the brighter children of a supervised class not only make better progress than the brighter members of equal ability of the unsupervised group but that they also make better progress than the slower children of the supervised group.

The experiment seems to show that the methods used in the direction of the supervised group had sufficient merit to lead to decided improvement of the supervised over the unsupervised group and to indicate further that the brighter members of the group were helped at least as successfully as the slower members.

Conclusions

1. Supervised study in English composition is eminently worth while when teachers are able to direct children's efforts intelligently. For this purpose they must know the technique of learning in this branch and apply such knowledge to the varying needs of the individual members of the class.

2. While a knowledge of the various general factors of study is worth while, at least to the teacher, she must not stop there. Methods of work imply a large number of specific knowledges, habits, and attitudes which differ in the several fields of study and which must be acquired through successful completion of specific tasks in each field. The pupil can acquire effective methods of work only by working with purpose and intelligence under a teacher who knows intimately the most successful methods of procedure.

3. As in all teaching an intimate knowledge of children's interest and experiences as well as capacities is absolutely essential in directing their work.

4. In the present stage of progress in the teaching of English composition the writer believes that teachers of this subject would be greatly benefited if they had access to a case book describing how the needs of the various types of children can be best met.

BIBLIOGRAPHY IN SUPERVISED STUDY

J. M. ALLEN, Experiments in Supervised Study. *School Review*, 25: 398-411. June, 1917.

FRED S. BREED, Measured Results of Supervised Study. *School Review*, 27: 186-204, March, 1919.

E. R. BRESLICH, Supervised Study As a Means of Providing Supplementary Individual Instruction. *Thirteenth Yearbook of the National Society for the Study of Education.*

E. R. BRESLICH, Teaching High School Pupils How to Study. *School Review*, XX: 505-515, October, 1912.

J. S. BROWN, Supervised Study in High Schools. *School and Home Education*, 24: 735-45, December, 1916.

A. W. BURR, Directed Study. *School Review*, 27: 90-100, February, 1919.

O. W. CALDWELL, Detroit Central High School Plan. *Popular Science Monthly*, March, 1913.

O. W. CALDWELL, The Influence of Prolonged and Carefully Directed Study. *N. E. A. Proceedings*, 1912, 691-700.

T. R. COLE, One Year of Supervised Study. *School Review*, 25: 331-36, May, 1917.

GRACE A. DUNN, Value of Supervised Study. *Teachers College Record*, Vol. 18: No. 5, November, 1917.

LIDA B. EARHART, *Teaching Children to Study.* Riverside Educational Monographs.

J. E. ERICKSON, The Results of Supervised Study in the Houghton, Mich., High School. *School Review*, 24: 752-58, December, 1916.

CHAS. E. FINCH, Junior High School Study Tests. *School Review*, 28: 220-226, March, 1920.

F. M. GILES, Investigation of Study Habits of High School Students. *School Review*, 22: 78, April, 1914.

C. H. HANDSCHIN, Individual Differences and Supervised Study. *Modern Language Journal*, Vol. III, No. 4, January, 1919.

G. L. HARRIS, Supervised Study in the University of Chicago High School. *School Review*, 24: 490-510, September, 1918.

H. C. HINES, Supervised Study in the Junior High Schools. *School and Society*, VI, No. 149, November 3, 1917.

KARL J. HOLZINGER, Periodical literature on supervised study during the last five years. *Elementary School Journal*, 20: 146-154, October, 1919.

JONES, LEARY AND QUISH, *Teaching Children to Study.* Macmillan, 1909.

L. I. LOVELAND, Supervised Study. *School Review*, 23: 489-90, September, 1915.

F. S. LUNT, Some Investigations of Habits of Study. *Journal of Educational Psychology*, June, 1910.

A. LAURA MCGREGOR, *Supervised Study in English.* Macmillan. Hall-Quest series.

F. M. MCMURRY, *How to Study and Teaching How to Study.* Houghton, Mifflin & Co.

A. S. MARTIN, Long School Day and Directed Study. *Education*, 39: 158-163, November, 1918.

E. D. MERRIMAN, Technique of Supervised Study. *School Review*, 26: 35-38, January, 1918.

J. H. MINNICK, An Experiment in the Supervised Study of Mathematics. *School Review*, 21: 670-675, December, 1913.

KATHERINE MORSE, Supervised Study. *English Journal*, 8: 355-62, June, 1919.

C. H. NIELSON, An Innovation in Supervised Study. *School Review*, 25: 220, March, 1917.

W. M. PROCTOR, Supervised Study on the Pacific Coast. *School and Society*, VI, No. 142, 326-8, September 15, 1917.

W. C. REAVIS, Factors that determine the habits of study in grade pupils. *Elementary School Teacher*, 12: 71-81, October, 1911.

W. C. REAVIS, Importance of Study Program for High School Pupils. *School Review*, 19: 398-405, June, 1911.

G. E. RICHARD, *Some aspects of supervised study*. Master's thesis, Chicago University.

A. C. ROBERTS, Supervised Study in the Everett High School. *School Review*, 24: 735-45, December, 1916.

MABEL SIMPSON, The divided-period plan of supervised study in American history. *N. E. A. Report*, 1919.

G. W. WILLET, *Supervised Study Technique*. Midland Schools, March, 1918.

SOME INDIVIDUAL DIFFICULTIES IN THE STUDY OF MUSIC

ESTHER ALLEN GAW, PH.D.
San Francisco State Normal School

One of the requirements of the Board of Education in California is that every teacher shall be able to teach the music of the grade to which she is assigned. She must, therefore, be able to sing the songs of the "Progressive Series," which is the prescribed textbook, and to read at sight readily enough to teach any other songs that may be needed. She must understand the technique of our music system, that is, the keys and their relationships and the staff notation.

In order to teach the staff, the keys, and the relationships of the syllables, the music department of the San Francisco State Normal School uses a series of bulletins which outline the minimum requirements. These are arranged in three units upon each one of which the student passes an examination. These units are as follows:

1. The student must be able to sing and write the scales of the tonic chord, in the keys of e-flat and f-major, getting the pitch from the pitch-pipe. These keys are used because they are within the staff and are pitched where the children sing.

2. The student must be able to sing at sight each measure of an exercise or song in 4-4, 3-4, or 6-8 time, with one sound to a beat, one sound to two beats, one sound to three beats, one sound to four beats, two equal sounds to one beat, and, in the case of 6-8 time, one sound to five or six beats using the keys and intervals studied in 1.

3. The student must be able to recognize and sing at sight skips of thirds, fourths, and fifths in the same two keys, and to apply them in the songs of Book 1, and also to pass a sight reading examination in an unknown song involving only these problems.

This is the outline of the course in music which must be mastered by any student before she is allowed to teach the simple primary songs.

In the fall of 1920 a survey of the students in the San Francisco State Normal School produced the fact that there were twenty-four who were conspicuously unsuccessful in their study of the music course after having carried it for three months. These students were not able to sing the songs alone either at sight when presented on the written page, or in true intonation. In order to determine whether or not the difficulty was one of inability to hear the tone differences which they could not produce, they were given the four sensory discrimination tests and the tonal memory test which have been recorded on the phonograph by Seashore.

The Seashore tests do not need a detailed description here. Suffice it to say that there is a test for pitch discrimination, one for intensity discrimination, one for the discrimination of duration of time intervals, one for the discrimination of the relative consonance of musical tones, and, finally, a test for immediate memory span for a series of nonsense musical tones.[1] These tests have been standardized and norms of percentile rank are published in the Manual.

To return to our students who are trying to learn to sing. The supervisor asks, "What is the cause of this inability to learn elementary music? Is it a general cause, or are there differences in their native capacity which make music so difficult for them to grasp; and how can I help each one most effectively?"

Before the tests were given, the twenty-four girls were sorted into three groups: (1) Six students of whom there was some hope that slowly and gradually they might come up to the required standard. These were called "hopeful." (2) Six students of whom the supervisor was in grave doubt, but of whom she still entertained a hope that by unusual diligence they might reach the standard. These were called "doubtful." (3) Twelve students who, the supervisor felt, gave unmistakable evidences of never being able to meet the required standard. These were called "hopeless." This ranking was done by the supervisor,

[1] Seashore, *Psychology of Musical Talent*, Silver, Burdett & Co. See also *Manual of Instructions and Interpretations for Measures of Musical Talent*, Columbia Graphophone Co., New York.

entirely independent of the psychologist, and purely upon her observation of the actual achievement of the students, but in the light of her teaching experience.

Table I shows the percentile ranks of the twenty-four students according to the Seashore tests. In percentile ranking the highest possible mark is 100, the lowest is 1, and the median is 50. An average of percentile ranks does not have any particular significance, since the tests are of different values in diagnosis, but it is desirable to have some general designation of the individual record as a whole. For the purpose of grouping I have called the records which on the average fall between 1 and 15 in rank, very inferior; those from 16 to 30 inferior; those from 31 to 47 fair; and those above 48 good.[2]

TABLE I.—PERCENTILE RANKS ACCORDING TO THE SEASHORE TESTS

	Student's Number	Pitch	Intensity	Time	Consonance	Memory
Hopeful	1	36	82	83	68	16
	2	87	74	32	6	60
	3	56	45	24	86	37
	4	87	22	54	46	30
	5	63	25	20	56	..
	6	29	25	32	56	13
Doubtful	7	10	58	61	56	30
	8	87	14	20	18	41
	9	29	19	73	1	5
	10	3	51	24	18	2
	11	1	4	2	78	3
	12	26	25	3	3	..
Hopeless	13	1	66	78	56	2
	14	9	87	28	36	34
	15	29	58	37	18	23
	16	13	66	16	46	12
	17	5	51	24	26	11
	18	1	5	3	78	3
	19	26	4	24	3	..
	20	1	19	9	36	3
	21	4	22	32	1	5
	22	1	10	28	..	3
	23	1	19	13	..	1
	24	3	4	4	1	1

[2] *A Survey of Musical Talent in the Des Moines Public Schools*, Child Welfare Research Station, Univ. of Iowa, 1920.

When it is understood that the rating hopeful, doubtful, and hopeless was independent of the sensory tests, and when the rank of the students in each category is studied, a distinct correlation is immediately perceived. This is more fully brought out in Table II. In this table the divisions good, fair, inferior, and very inferior are based upon the general results of the five

TABLE II.—RELATION BETWEEN SUPERVISORS' JUDGMENTS AND RATINGS ACCORDING TO THE SEASHORE TESTS

Supervisor's Classification	Classification According to Seashore Tests			
	Very Inferior	Inferior	Fair	Good
Hopeful	0	0	2	4
Doubtful	1	3	2	0
Hopeless	6	3	3	0

measurements of each individual. The divisions hopeful, doubtful, and hopeless are the empirical classifications of the supervisor. The figures in the spaces indicate the number of individuals under each of the double classifications. It is apparent that those who are hopeful are the highest in their sensory capacity, that those who are doubtful fall a little lower, and that those who are hopeless are undoubtedly the lowest in the sensory measurements.

I shall describe in detail a few of the cases. Each one is interesting and significant, but I shall limit my descriptions to a few of the more dramatic ones. I shall refer to them by the numbers used in Table I.

The supervisor's comment on No. 1 in September was that she had no experience and was very timid. It was this timidity that made me feel that the results of the first group tests were probably unreliable, and I tested her again with the results given in the table. Three months later the supervisor reported that while she was not especially strong in music and was very lacking in self-confidence, she was slowly learning the required songs and would be able to teach the simple songs at least. The student herself reports that when she learns the songs with the syllables she cannot then transfer the words to the melody and, therefore, is very slow in learning the material. The process demanded

by the teachers (that of first learning the syllables) which is a help to many, is a hindrance to her. Her motor coordinations are persistently inflexible with this resultant difficulty. She also has a distinct sensory inferiority, since she does not remember tones, as such, as well as the majority of normal people do. These two weaknesses, in connection with a lack of self-confidence, will make her slow, but there is no reason why she cannot ultimately learn the music required.

No. 2 is a case of another timid soul who does not show her true capacity at first. In September the supervisor reported that she could not sing a melody alone. Three months later she reported that she was singing with great timidity and getting her work very slowly, but that she was able to sing the simple primary songs with an accurate voice of pleasing quality. It is not a sensory difficulty, then, as the tests show, nor a motor difficulty, as the accuracy of her singing shows, but a temperamental difficulty; and she is, therefore, properly classified as "hopeful."

No. 8 was found to be fair in her sensory capacity. The supervisor said: "Her voice flats; she cannot remain true." Three months later, although the student had been working faithfully at her music, the same report was made. In the supervisor's words: "Her voice flats very markedly; she goes quite off and it is a surprise to her that she isn't doing it correctly. She only approximates a tune. She can read and is trying hard to finish her music. If she could retain the pitch, she would be all right." No. 8 is one in whom the sensory capacity is not the important weakness; her difficulty is in motor capacity. There are people who can see a ball and who can hold a tennis racquet, but who have great difficulty in learning to hit the one with the other. In the same way there are those who hear the pitch of tones and who can use their voices for ordinary purposes, but who cannot hit the pitch of the tones accurately with their voices. No. 8 is an illustration of this kind of motor difficulty. She will have to take very intensive practice in gauging the pitch distances which her voice must control; and by such drill alone will she be able to improve the accuracy of her singing.

No. 10 was reported in September as having no musical experience. Three months later the supervisor said of her, "She appears to have no conception of what reading means. She doesn't look at the right place in the book. She can imitate tones but seems to have no power to concentrate on the problem of reading." With her inferior rank in the sensory tests she probably cannot remember the tones for any length of time. She may have no particular weakness in motor capacity, but if, in addition to her inability to learn to read the notes, she is further handicapped by inferiority in intelligence, she will probably never be able to do even the minimum in music.

No. 15 was reported as being weak in music, but the supervisor was not sure whether this was due to lack of previous experience or to natural incapacity. Three months later the supervisor had found that No. 15 could not yet sing a melody and had very little conception as to whether or not she was singing correctly. She displays fair sensory capacity but is probably lacking in motor control and coordination. She is very intelligent and could undoubtedly improve her motor attack to some extent.

No. 17 was simply reported as being weak in music. Later the supervisor noted that if she worked on a tune long enough, she got it but not with perfect accuracy. Another supervisor asked her in my presence to sing a tiny melody of four tones not more than a step apart. She made three attempts but went in the wrong direction on each trial. The tests show that she is distinctly inferior in her sensory capacity, and she is also evidently inferior in her motor capacity. Whatever she learns to sing will be acquired slowly and with great difficulty and will never be accurately retained. She loves, however, to listen to music and works diligently at it, although she knows that she is accomplishing very little.

No. 22 was reported as having no previous knowledge of music. Three months later the supervisor's report says, "She sings down when the notes go up, and vice versa. Time values she also cannot comprehend, although she has had special training in a small class. A song that she once learns, largely by

hearing it, she can sing." Here is a case of very inferior sensory capacity linked with inferior intelligence which means that she will learn even the minimum amount of music only by the greatest expenditure of effort on her part and specially adapted training on the part of her teachers.

From these briefly described cases some general observations may be made. Difficulties in learning music are of at least two distinct kinds; there are sensory weaknesses and there are motor weaknesses. Each may exist in an individual by itself or in connection with the other. Either one alone puts the possessor into the class of those who learn music slowly, painstakingly, and —often in the end after much effort—ineffectively. If to these weaknesses is added slowness in general intelligence the situation becomes even more difficult. Those who are found to be of either or all of these three classes must be trained individually by methods specially adapted to the particular case.

At the San Francisco State Normal School an effort is being made to give special training to these students. The work is yet in the experimental stage, so that I cannot describe any final results. The students who were graded as hopeless have been grouped together into a special class, where they are taught to sing some simple primary songs together by note. These they sing with haziness of intonation but with a rather surprising approximation to the pitch. When each student sings alone, she is not at all sure of herself. The quality or timbre of their singing is very bad, and apparently they are very vague in a conscious conception of what they are doing. By this I mean that they are not accurate in the beginning and ending of phrases or in the actual relationship of pitches. They rely upon their teacher and upon one another. I felt when I heard them sing in a group and individually as I do about a little child who is learning to walk. I admire his tottering efforts when I consider that he is a little child. But if I consider his achievement as walking I must admit that it is not a very successful performance. So I feel about these students who are deficient on the sensory or motor sides of their inborn equipment for music; it is surprising

that they can sing at all, propping themselves up on each other and leaning on their teacher, but as real singers they are very bad.

The results of this study, reinforced by previous observations of similar cases, lead me to the belief that those who are weak or inferior in hearing the characteristics of musical tones will be weak and slow in reproducing them. It would undoubtedly be more economical of effort, both on the part of the student and of the teacher, if they could be released from the drudgery of trying to learn something which is so difficult, and could turn their energy and talents to a field where achievement would be more notable. But, since under the present circumstances we must train them as well as we can, we must reinforce their auditory-motor weaknesses by artificial associations. For instance, the tonoscope may be used to give a visual correction of faulty intonation, or an artificial association of tones with colors may be deliberately used to improve tonal memory. Various synesthesias, according to the needs and capacities of each individual, may be worked out as crutches for the lame ducks in music. This immediate problem of forming artificial associations is a very interesting one on which I hope to be able to report later.

INTELLIGENCE TESTS AND THE CLASSROOM TEACHER

ARTHUR W. KALLOM
Assistant Director of Educational Measurement, Boston, Massachusetts

Tests, either intelligence or achievement, may be given for three reasons: (1) for the purpose of collecting data; (2) for the purpose of satisfying the curiosity of the teacher; (3) for the purpose of bringing a definite constructive help to the teacher in guiding and aiding her in dealing with some of the problems with which she is confronted.

Up to the present, much of the work of educational measurement has been concerned with the first step. This is important. It has made it possible to establish standards, it has confirmed the theories that the educational product is variable and not constant, it has emphasized and placed on a more or less scientific basis, perhaps, the greatest educational fact of the past generation, namely, individual difference. There is still need for a great deal of work to be done along this line. This research work is the work of universities, foundations, and similar bodies and not the work of city departments, which are supported by public money. Their research work, *per se*, is a by-product.

It is a good sign to see the teacher's curiosity aroused to know what her class will do with certain tests. Perhaps she has a suspicion that *her* class will not show marked signs of individuality because *her* class was a selected group and because children of about the same grade of ability were put together in her room. Therefore she is curious to see if some of the theories of the savants can not be exploded and is much surprised to find that there is, perhaps, as great variation in her room as in another room not so carefully selected.

The teacher who is curious to know what will happen when a test is given should be greatly encouraged, because out of the "knowing" is going to come "doing." Out of the knowledge of where each pupil stands is going to come a better directed effort

to help those who need help, to drill those who need drilling, and to guide those who need guiding. The efficient teacher is the one who is accepting all the help which tests can give in promoting, drilling, and guiding. Such a teacher makes it justifiable for a city school system to expend public money for a department of research. Directly or indirectly such a department makes it possible for the taxpayer's son and daughter to be better and more efficiently taught than before; better taught because both teacher and pupil have a definite aim toward which to press, and more efficiently taught because the teacher with a more direct knowledge of what the child is able to do can save valuable time.

Intelligence Tests

During the past few years intelligence tests have held a prominent place in the testing program of departments of research and of all up-to-date superintendents. Such tests are fraught with a good deal of difficulty unless one is careful how the results are used and unless the one in charge realizes the limitations of such a test. Two cases will illustrate what I mean.

In a certain city one of the well-known group intelligence tests was used and a certain pupil scored low. The score was so low that the pupil was judged feeble-minded by the principal, and the parents were asked to withdraw the pupil from the school. The parents objected, and when he was tested by a competent person, using the Binet-Simon scale, the boy showed an intelligence quotient of 123.

The second case was a girl in one of our Boston schools who tested low with the National Intelligence Test but who was judged one of the brightest children in the school by the teachers. The possibility of such a condition was admitted, but an individual test was given. The girl showed an I. Q. of 115. It was found upon investigation that, on the morning the group tests were given the girl had come from home after witnessing a brawl between her father and mother in which she was obliged to take some part. She had had no breakfast and had taken the test within half an hour after arriving at school.

The giving of a group intelligence test in a public school system by the superintendent or director of research, may have one or more of a number of possible objects in view. In Boston there have been four main reasons for giving such tests: (1) to guide teachers in making promotions; (2) to guide pupils in their choice of courses; (3) to advance demoted pupils; and (4) to guide teachers in dealing with individuals who are failing in their work.

Use of Intelligence Tests in Making Promotions

In March, 1921, 45 districts were tested with the National Group Intelligence Tests. In spite of the request of most of the principals that all grades should be included, only Grades VI, VII, and VIII were tested, owing to the expense and inability to handle results.

Grade VI was tested in order that the results might be used in the promotion of children to intermediate or junior high schools. Grade VII was tested with the expectation that the eighth grade would be better organized for promotion to high school the succeeding year. Grade VIII was tested that teachers might have the data to guide them in promoting to high school. The succeeding fall the data collected from Grade VIII were also used by the high-school principals and teachers.

In promoting a pupil, two factors were taken into consideration—his test score, and his teacher's mark. If the two agreed, there was no particular difficulty in settling the question of promotion. If the individual's test score was high and the teacher's mark low, he was promoted on the basis of his test score. If the teacher's mark was high and the test score low, he was promoted on the basis of the teacher's mark. In either case, it will be noted, the pupil was given the benefit of the doubt. These pupils, therefore, became marked pupils and have been problems for study during the year. How far children who were promoted on teachers' marks have succeeded will not be known until the end of the year.

Intelligence Tests Aid in Guiding Pupils to Choice of Courses

Children entering the seventh grade in any Boston intermediate school, may choose to take a modern foreign language, either French, Spanish, Italian, or German. Each school may have only one foreign language, however. Because this modern foreign language is an extra subject, only those children are allowed to take it, whose marks are equivalent to an average of *B*.

Of 51 children in one school who were tested in the sixth grade and placed in French classes, 25 percent had a score of less than 109, the median of the sixth grade. Of this 25 percent, 39 percent made good, but in every instance effort and conduct were excellent. Of the 75 percent who were above the median, only 2 persons or 4 percent failed. One of these had an effort of *E* and a conduct of *D*. This school is typical of all intermediate schools.

Therefore, for a pupil who scores below the median the chances of success with a program which requires an extra amount of work are about two chances out of five, and then only because effort and conduct are excellent. When the pupil's score reaches 120, the chances of success amount almost to certainty.

When children enter high school they are required to select a definite course which may lead to a college, technical school, normal school, or business. In many cases children desire to take a college course, who are totally unfitted either to profit by such a course or to handle the necessary preparatory work. Such children are not warranted in spending their time either in high school preparing for college, or in college, providing they barely reach the minimum requirements for entrance.

Making a careful study of the relation between the score of an intelligence test and the marks obtained by an individual in his high-school work, will enable us to prophesy the chances which the individual may have to complete his high-school work successfully.

Table I shows that Pupil No. 1 is 14 years 8 months old. In the Binet-Simon Test he showed an intelligence quotient of 99.

In the Terman Group Test he received a score of 104; and in the Army Alpha Test a score of 107. The marks in his elementary and high-school subjects are in his major subjects.

Number 11 in Table I is a boy who came to one of our high schools at the age of 15. He entered from the Lyman School, which takes incorrigible cases sent by the courts. He has a bad temper and does not get along with other boys. He gave acceptable definitions for 45 words in the first column of the vocabulary used in the Stanford Revision of the Binet-Simon Test. He answered all but one question used in the Superior Adult test

TABLE I.—COMPARISON OF INTELLIGENCE SCORES AND TEACHERS' MARKS

Pupil	Age Sept. 1	I. Q.	Nat. Intel. Test	Terman Test	Army Alpha	Grammar School Marks	High-School Marks
1	14–8	99	...	104	107	B A B	C D C E
2	15–6	82	...	76	85	C C D	C D C E
3	13–1	107	141	68	95	C C C	C C C D
4	15–4	87	...	59	65	B B B	D C D C
5	13–10	87	117	93	95	C C B	B D D B
6	14–6	81	100	79	91	D C D	C D C E
7	14–7	88	...	38	62	C C E	E E D D
8	15–10	88	...	91	91	.....	D C D C
9	14–2	100	118	100	92	C C C	B B B B
10	13–1	97	...	59	...	C C D	D D C C
11	15–5	118	...	165	136	.....	B B A B
12	16–6	61	...	31	...	C C D	E E E E
13	14–7	81	110	69	59	C C D	C D D C
14	14–11	93	...	60	82	E E D	D C C D

satisfactorily. He is now taking two subjects in addition to his regular work and is planning to take trigonometry at an evening school. After an hour's talk (and he talked very freely), it is evident that the only trouble with the boy is the fact that he never had enough to do in school. He was a boy who could have done considerably more than was required but his spare time was used in getting into trouble.

A certain boy, A. B., in one of the high schools had been taking an industrial course for three years. His work was eminently satisfactory. An older brother, however, had graduated and entered technical school. This had created a desire in A. B.

to attend technical school also. He had previously tried the technical course and failed. The mother requested that the boy be tested to find out whether he was capable of doing the work of a technical school. The boy tested slightly above normal. After a long talk in which the boy agreed that he would rather be a good mechanic than a poor or fair civil engineer, it was decided that he should re-enter the mechanical course. At the last report the boy is happy and contented, doing a thing which he can do well.

A second case in the same school, a boy who remained there because his father so insisted, had an I. Q. of 79. He had failed in freshman classes for two years and was failing the first year in the sophomore class. Since giving the first test the boy has been withdrawn. The practice, however, of withdrawing a boy who is not keeping up, even though he may not be able to do so, is not a solution of the matter. The boy has a right to as much of an education as can be given. The state is insisting on more and more.

Determining mental ability for purposes of differentiating between those fitted to do the work now required at any stage of the educational program, and those unfitted to do the work, means the organizing of a suitable program for those who find the present curriculum ill-adapted to their particular needs. In determining this mental ability, an important question arises regarding the extent to which it is wise to go in our differentiation.

Intelligence Tests Aid in Advancing Demoted Pupils

Many children are obliged to repeat a year on the basis of teachers' marks. It is known that teachers' marks are to a large extent unreliable. This results sometimes, in requiring children to repeat a year's work, who are mentally capable of doing the work of the next grade.

In testing certain classes early in September, 1920, the poorest class in a certain building was easily identified, and it was interesting to find that this poorest class was composed of seventh-grade pupils who had not been promoted. Of the 33 children

in this class, however, seven were found to be mentally superior. On the basis of the intelligence test, these seven children were granted permission to go into the eighth grade on trial, providing they showed good effort and good conduct. Four out of the seven made good.

Intelligence Tests Guide Teachers in Dealing with Individuals

Children are constantly failing in their school work, and many times the teacher is unable to tell, either certainly or quickly, the real cause of their failure. Even though the intelligence test has not yet become a perfect instrument, it does give an indication of what ought to be expected of a given individual. In many cases the pupil is recognized as a good pupil, but is not giving the kind of work of which he is capable.

C. D. was a boy, who had been judged by teachers to be of good material, but lazy. Teachers had always said, "*I think* he has good ability." The father said, "*I believe* he has good ability." The boy insisted he was doing all he could. His score in the intelligence test was the highest in the district. The statement of teachers and the father now changed from "I think" and "I believe" to "I know." The boy recognized that he was caught and changed from a fair scholar to a superior one.

E. F. was a girl who gave a great deal of trouble. Her father was a constant visitor to the school. He insisted she was a bright girl. The teachers constantly insisted that she had never shown any signs of a high degree of intelligence. The test showed her among the best. The problem now is, "How can the teachers get from this pupil the work she is capable of giving?"

Intelligence Test in an Actual Schoolroom

Figure 1 shows the seating plan of an actual sixth-grade room. The school is situated in a Jewish quarter, and all but two or three children in the room are of Jewish origin. Most of them were born in the United States, and have had the advantages of the city schools up to the sixth grade. In most

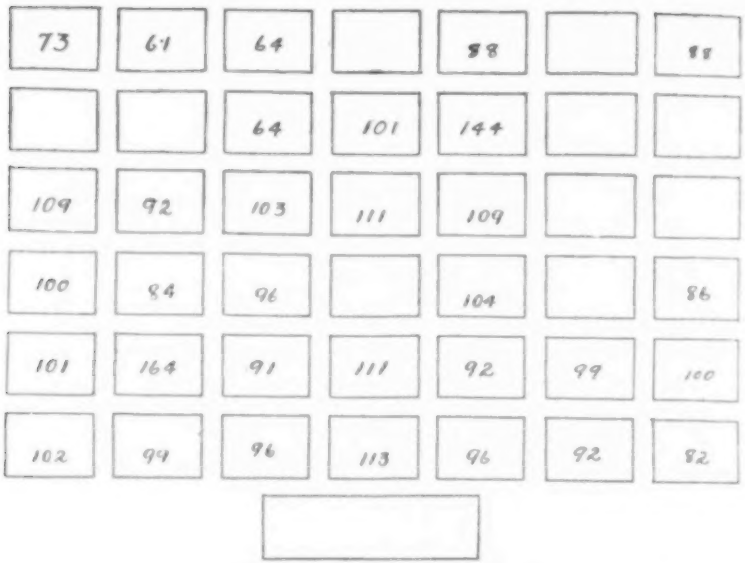

FIGURE 1.—SIXTH GRADE SEATING PLAN SHOWING PUPILS' INTELLIGENCE QUOTIENTS

cases, however, the language of the home is Yiddish, because the mother rarely speaks English. The numbers placed in the small rectangles represent intelligence quotients based on the Stanford Revision, all children having been tested either by the author or under his direction. It will be noted that the intelligence quotients range from 61 to 164.

The teacher of this room is, therefore, obliged to face each day children ranging all the way from superior mentality to a mental condition which borders on, if it is not actually equal to, that of a high-grade moron. She is expected by principals and other supervisory officers to prepare these individuals for promotion to Grade VII, with the expectation that after they have been promoted to Grade VII they will be able to do the work of that grade. The intelligence test often shows the utter impossibility of any teacher's doing the work that is expected of her.

Figure 2 shows the same schoolroom tested with the Thorndike-McCall Reading Test. The numbers in the small rectangles

indicate the reading quotients of these same individuals. Owing to absences there are one or two discrepancies between the individuals of Figure 1 and Figure 2.

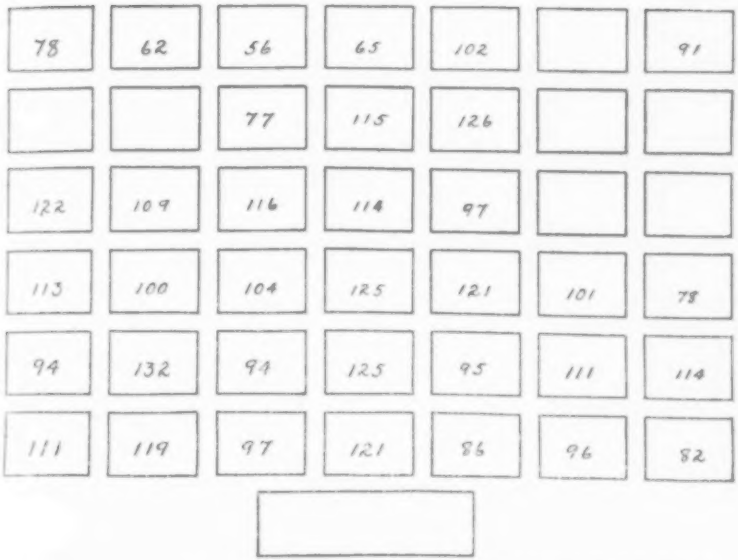

FIGURE 2.—THE SAME SEATING PLAN, SHOWING READING QUOTIENTS

These reading quotients indicate the same variation between the pupils as was shown by the intelligence quotients in Figure 1. A close study of the two quotients will show also a fairly close correlation between them.

Intelligence Tests an Aid in Organizing Classes on a Basis of Mental Ability

It seems to the writer an open question how far the intelligence quotient can be determined from the results of the group test. The factors determining the results in the intelligence test are so variable, that it does not seem possible to make them constant enough at the present time to determine a reliable intelli-

gence quotient. It also becomes a question whether such a fine discrimination as the intelligence quotient gives is necessary in grouping children in accordance with their mental ability.

It has been the practice during the past year to group the scores on an intelligence test into a distribution based on scores of the grade, and then to divide that distribution into quartiles. It is then possible to place each pupil's score into its proper quartile. This has been called the grade score. Following this a distribution has been made of all the scores on the basis of age regardless of the grade. This has also been divided into quartiles. The placing of each score in its proper age quartile has been called the age score.

Figure 3 represents the report to the master of the room showing the name of each pupil, his sex, age, total score, and quartile standing as to grade score, age score, and age. Thus any pupil who is in the fourth quarter in his grade and age score

NATIONAL INTELLIGENCE TEST SCALE A—FORM 1

.....DISTRICT

March 16, 1921

Grade VI Teacher.....

Median Score (as given by National Research Council), 112

Boston's Median Score.....109

Name	Sex	Age	Pupil's Score	Quartile Standing		
				Grade Score	Age Score	Age
	G	13-9	81	4	4	1
	G	11-6	121	2	2	3
	B	11-6	87	4	4	3
	G	11-2	123	1	2	4
	G	11-3	106	3	3	4
	B	11-0	66	4	4	4
	B	11-6	95	3	4	3
	B	11-6	111	2	3	3
	B	13-5	91	4	4	1
	G	10-8	112	2	3	4
	G	10-10	110	2	3	4
	B	12-0	133	1	2	2
	G	11-0	121	2	2	4
	G	11-3	144	1	1	4
	B	11-1	153	1	1	4

FIGURE 3.—REPORT ON THE INTELLIGENCE OF INDIVIDUAL PUPILS

is a poor pupil. Any pupil in the first quarter in his grade and age score is a superior pupil. Such a method makes it possible to divide groups of pupils into four classes based on their general intelligence. It brings all the superior pupils into one group, and those who are mentally inferior into another group. It obviates such a situation as is shown in Figure 1, where the general intelligence of the class is very widely distributed.

It is also probable that at present we cannot hope to divide any large group of pupils into more than four groups. It certainly would be unwise to attempt to divide pupils into groups having the same intelligence quotient; even if it were possible, the number of divisions involved would necessarily be so large as to be unwieldly.

Conclusions

1. Intelligence tests should be given for the purpose of bringing to the teacher a definite constructive help to guide her in solving the problems with which she is confronted.

2. Educational research in a public school system is only justifiable when it directly or indirectly aids the teacher in solving her problems.

3. Intelligence tests may be used: first, to guide teachers in making promotions; second, to guide pupils in the choice of courses; third, to advance demoted pupils; and fourth, to guide teachers in dealing with individuals failing in their work.

4. Pupils should be divided within the grade on the basis of their mental ability as determined by group intelligence tests, supplemented by individual tests.

5. Wherever the number of pupils is large enough to warrant it, pupils should be divided into four groups.

SCALES FOR MEASURING RESULTS OF PHYSICS TEACHING

HAROLD L. CAMP
University of Illinois

Those who are familiar with the teaching of high-school physics realize the need of scientific investigation to aid in solving a number of vital questions. The following are some of the most significant of these. To get the results sought, how much time should be spent in performing laboratory experiments? Do the results of laboratory work justify the expense involved in furnishing several sets of equipment? How should the time for the study of high-school physics be distributed as to: work in the laboratory, learning the fundamental principles, solving problems, recitations, and observation of class-room demonstrations? How may the course of study be revised to meet more fully the needs of the majority?

In attempting to solve some of these problems one is confronted with the need of suitable tests or scales with which to measure physics ability or achievement. If we have such measuring instruments we may by varying the time devoted to laboratory experiments ascertain the resulting degrees of progress. Similarly, the relative advantages of single and duplicate sets of apparatus, of different combinations of laboratory work, observation, problem solving, etc., and of various other devices and procedures may be determined through the achievement of the pupils.

Few outstanding contributions have been made to the field of high-school physics testing. Starch[1] arranged 75 mutilated sentences involving 102 fundamental principles treated by five most commonly used textbooks; but he did not give the relative difficulties of these 75 test exercises. Chapman[2] reported the per-

[1] Starch, Daniel, *Educational Measurements*. New York: The Macmillan Company, 1916. pp. 188ff.

[2] Chapman, J. C. "The Measurement of Physics Information," *School Review*, 27: 748-56, December, 1919.

cent of pupils in Cleveland high schools who correctly solved each of 30 or more exercises devised and selected by a group of physics supervisors and teachers. E. L. Thorndike has devised several physics tests for college students but none for high-school students.

Because of the need for additional material to be used in testing high-school physics achievement, the writer has derived "Scales for Measuring Results of Physics Teaching" which have been published by the State University of Iowa in its *Studies in Education*, volume II, number 2.

On the basis of the aims of science teaching in the high schools of the North Central Association as reported by H. A. Greene,[3] the primary objects of physics teaching were taken to be: (1) to give knowledge of natural phenomena, (2) to serve as a guide for daily life, (3) to impart a scientific attitude, and (4) to prepare for college entrance. In formulating and selecting test exercises for the physics scales, the writer has assumed that the two most important aims are: (a) knowledge of the fundamental principles, and (b) ability to put such knowledge to use in solving the problems one meets in ordinary life.

The task was, then, to formulate, select, and evaluate exercises, the correct answering of which would involve one or both of these types of physics ability. It was also necessary to have exercises whose difficulties varied by easy stages from rather easy to extremely difficult. The test exercises finally selected are of two general types, factual and problematic; and they are designed to measure ability in (1) mechanics, (2) heat, (3) electricity and magnetism. The following exercises are typical:

> "What is the common name of the instrument used to measure the pressure of the atmosphere?"
>
> "What is the common name of the time-rate of change of velocity?"
>
> "In the ordinary electric light bulb there is little or no air. When a bulb is broken will the glass start moving toward the center or away from the center of the bulb?

[3] Greene, H. A., *Status of the Sciences in North Central High Schools*, State University of Iowa, Thesis, (1917).

"What is the density of a block of wood that floats ¾ under water?"

"What is the possible buoyant force of 78 pounds of cork whose density is 15.6 when water weighs 62.4 pounds per cubic foot?"

In selecting the exercises the following requirements were used as a guide: (1) The test exercise must apply specifically to the field of knowledge to which the test is related. (2) It must be taken from an essential part of the subject, a part which is contributory to understanding or summarizing a general theory or principle; in other words, it should be such that a correct answer would indicate that a distinct meaning had been grasped. (3) It must be so worded as not to be ambiguous. (4) It must be so framed that it can be answered very concisely, usually by a single word, or at most by a short phrase. (5) It must be such that the short answer given in the test includes all possible correct answers.

The exercises were arranged in 47 different preliminary tests of nine or ten exercises each. The correct solutions of these exercises were attempted by approximately 3500 boys and girls in 129 of the best high schools in Iowa. Each test was distributed over the state so that the pupil-reactions on the exercises included in it would be representative of the state as a whole. On the average each exercise was attempted by 99 different pupils (minimum 60); and each test was given from eight to ten weeks after the work in the field to which it applied had been finished.

The Probable Error method was used in evaluating the exercises. The results, shown by a statistical treatment of the data gathered, justify the fundamental assumption of this method, namely that the distribution of physics abilities conforms approximately to the Normal Probability Curve of Distribution. An exercise so easy that only 1 in 10,000 would fail to solve it correctly was assumed as an exercise of zero-difficulty. Its position on a linear scale of difficulty is at —6 P. E. or about (—4 sigma); and all the test exercises finally used in the scales were assigned

values with reference to this location as a point of departure, or zero point.

As a result of this investigation, three scales have been developed including 289 exercises in all. These scales are distinctly separate. As has been stated, they include exercises to measure ability in the three fields of Mechanics, Heat, and Electricity and Magnetism. In each case about half of the exercises are designed to measure "knowledge of fundamental principles" and half to measure "ability to make useful application of such knowledge."

For each exercise the average amount of time used by those who correctly solved it has been determined so that from the scales we may select exercises to form both rate tests and power tests for any one or more of the three phases of physics. If desirable, one can form a test from either type or from both types of tests exercises, i.e., factual or problematic. The author has formed ten tests, (including both types of exercises), four forms for mechanics, three for heat, and three for Electricity and Magnetism,—which will be available soon from the Extension Division of the State University of Iowa. The general character of the tests may be inferred from the following exercises which are in each case the first or easiest five of Form I in Mechanics, in Heat, and in Electricity and Magnetism.

MECHANICS

1. What is the common name of the instrument used to measure the pressure of the atmosphere? (value 4.3)
2. If a 50 pound ball falls 100 feet and all its energy is transformed into work, how much work will be done? (value 5.3)
3. In the ordinary electric light bulb there is little or no air. When a bulb is broken will the glass start moving toward the center or away from the center of the bulb? (value 6.0)
4. What is the density of ice when 100 cubic centimeters weigh 92 grams? (value 6.8)
5. What is the efficiency of a machine when a force of 50 pounds acting through a distance of 30 feet lifts 200 pounds 6 feet? (value 7.3)

HEAT

1. What is the common name of the process by which the sun transmits heat to the earth? (value 4.8)

2. From the following select a good conductor of heat: wood, stone, air, copper, water, vacuum, iron. (value 6.3)
3. What is the common name for the action which takes place when an iron ball is warmed from 10 degrees C. to 20 degrees C.? (value 6.9)
4. At ordinary atmospheric pressure ice melts at 0 degrees C.; to make ice melt at —2 degrees C. would we increase or decrease the pressure? (value 7.8)
5. What is the common name of the process by which heat is carried from one place to another by a moving fluid? (value 8.5)

ELECTRICITY AND MAGNETISM

1. From the following select one which is a poor conductor of electricity: graphite, glass, hard rubber, pure water, paper. (value 3.6)
2. What is the common name of the electro-motive force needed to drive a current of one ampere through a resistance of one ohm? (value 5.7)
3. Will positive electrification attract negative electrification? (value 6.3)
4. With what kind of a cell should the circuit be kept open when not in use? (value 7.8)
5. What kind of positive and negative static electricity is formed on a conductor in the neighborhood of a charge? (value 8.7)

Table I shows the tentative norms on Form I of each test for boys and girls in Iowa high schools.

TABLE I.—TENTATIVE NORMS (FORM I)

	Median Score			Inter-quartile Range		
	Boys	Girls	Total	Boys	Girls	Total*
Mechanics	39.5	30.6	33.5	24.6—53.3	18.4—44.3	21.0—49.0
Heat	44.4	40.1	41.9	29.9—58.1	25.8—53.1	27.4—55.1
Electricity and Magnetism	46.3	36.5	39.9	33.2—61.8	25.5—49.3	28.0—54.6

* Total refers to both boys and girls.

The three scales and the tests available therefrom, when properly selected, administered, and scored, should prove to be of considerable value: (1) to the experimenter, as a means of comparing the results of two or more methods of teaching the same subject matter; (2) to the college or university, as a part of

entrance examinations; (3) to those who make school surveys; (4) to supervisors, in comparing the progress and attainment of different classes or schools and the results of different methods of teaching or of different classifications of pupils, in determining the amount of time to devote to different phases of the subject, and in determining what the course of study shall include; (5) to teachers in checking the results of teaching, as a guide in making the instruction more practical and valuable for all in a given group, and in determining the specific needs of individuals.

CONVENIENCE AND UNIFORMITY IN REPORTING NORMS FOR SCHOOL TESTS

J. Crosby Chapman
Yale University

The desirability of securing some uniform procedure for scaling the achievements of pupils in school tests is evidenced by the recent attempts and suggestions which have been made by several authors. Reference may be made to the work of McCall,[1] Pintner,[2] Buckingham,[3] and Buckingham and Monroe.[4] These very recent papers indicate the extent to which trained workers in this field realize the necessity for some definite conventions. For the larger number of individuals without much training, who are using these tests in practical work, there is even more necessity for definite agreement. An association such as this journal represents could hardly fulfil its function better than in furthering this attempt to secure uniformity and order, in a situation which as it now exists is chaotic.

At present a score of twelve in one test means the same level of achievement as fifteen, thirty-five, or seventy-one in another test. Even after we have transmuted from the crude scores to equivalent scores, there is again no uniformity. A score of fifty may mean the median achievement of unselected twelve-year-olds, as in McCall's plan, or the median achievement of an unselected group of any age as in Pintner's procedure, or the achievement of a particular mental age as in the method of Monroe and Buckingham.

We must work for an agreement whereby after the transition has been made from the crude score to the equivalent score, the latter will have an obvious meaning. This meaning must be

[1] McCall, William A. "A uniform method of scale construction," *Teachers College Record*, 22: 31-51, January, 1921.

[2] Pintner, Rudolf and Marshall, Helen. "A combined mental-educational survey," *Journal of Educational Psychology*, 12: 32-43, January, 1921.

[3] Buckingham, B. R. "Suggestions for procedure following a testing program," *Journal of Educational Research*, 2: 787-801, December, 1920.

[4] Buckingham, B. R. and Monroe, Walter S. *Illinois Examination, Teachers Handbook*, 1920.

simple, and above all it must be expressed in terms immediately applicable to the practical uses to which tests are put by schoolmen. The present almost complete lack of uniformity in reporting standards is unquestionably complicating and in many cases preventing the use of educational scales in the schools.

Binet and Terman have fulfilled the above requirements in the field of mental measurement. Perhaps, as is indicated later, their method of reporting mental ages has sacrificed too much to the desire to be readily interpretable by the layman, but they have at least used a very simple unit of measurement. Possibly the wide use of intelligence scales is partly explained by the obvious interpretation of their results.[5] It is certainly not too much to ask, that in the near future a somewhat similar uniform procedure may be adopted in the field of educational measurements.

McCall makes the plea that all tests be standardized and the equivalent scores be expressed in terms of an adapted sigma distribution of unselected twelve-year-olds. Involving as it does a continuous scale; many advantages are at once derived from such a procedure; these are well summarized in his article. In discussing some of the weaknesses of his proposal, these merits must not be overlooked. The following disadvantages must, however, be given careful consideration. Some of them, in the opinion of the writer, are sufficient to prevent his procedure from being adopted in school practice.

(1) It is impossible to scale even the moderately low performances of the lower grades, and the high performances of the upper grades. This is because the data for twelve-year-old pupils become so extremely attenuated or non-existent as to make any scale score unreliable, or impossible to determine.

(2) McCall's basic assumption is very much open to dispute. This is that however irregular a distribution of the twelve-year-old population may be, the distribution of all the other age populations is exactly, or within reasonable limits, the same. While

[5] At least it appears obvious on the surface; this obviousness disappears on closer analysis.

such an assumption may well be true for tests which are very carefully constructed, and for tests which measure some skill such as reading ability by material which is easy for the lowest age population considered, it is more than doubtful whether the assumption holds for tests which are less carefully assembled, and for tests which measure informational material rather than some basic skill of steady growth.

(3) There will be a natural tendency on the part of the casual user of tests so standardized to assume, for a particular group, other than twelve-year-olds, that forty on the scale for one test is really the equivalent for that group of forty on the scale for any other test. Great practical advantages would be derived from any scheme so constructed that, for any particular group considered, a score of forty would mean the same level, with respect to that group, in any test, just as in McCall's scale a score of fifty or any other score means the same, for the twelve-year-old population, in every test. Owing, however, to the differences in overlapping of any two groups in different tests a score of forty on his scale in one test would be attained by x percent of a particular population, other than twelve-year-old, while the same score in another test would be attained by a quite different percent. Thus while a scale score of t units means the same, irrespective of the test, for twelve-year-olds, it would not have the same meaning for any other population.

(4) There seems something incongruous in standardizing tests, covering material taught in Grade VIII or VII, on a population of which a considerable portion has not been subject to instruction. This point will be returned to later.

(5) There is the less academic but more practical question as to whether the school people who give the tests are not more interested in thinking in terms of grades, that is in placing a child in a particular group, than in discerning his position on a scale calibrated with an unselected group of twelve-year-old pupils. Usually the question is one of promotion, or transfer to some other group. Where the thinking is done in terms of grades there is a great advantage in having the median and other per-

formances of each grade in some constant term. While unquestionably as far as grade seven is concerned the median performance on the McCall scale will be approximately fifty, for all the other grades, the standards and achievements of the various levels will have to be expressed in units which will be by no means self-evident. Thus a score of thirty when made by a fifth-grade child will not carry its meaning on its face; it will only, after reference to further tables, tell us how he stands in his own grade.

For purposes of standardizing new tests, there are great advantages in McCall's method, but for convenience of use in the schools, it would seem that a method centering around grades has very distinct advantages. Of course McCall's scale could again be turned into levels of achievement at the various grades. For the schoolman, however, this is a gratuitous intermediate process.

Pintner has employed almost the same method as McCall, but instead of using a single age group, he uses all the unselected age and grade groups. He also is guilty of pressing his data beyond the point at which reliability can be obtained.

Furthermore, while the small word "unselected" can easily be inserted as a requirement in the standardization process, an unselected age group is irksome to secure, whereas an unselected grade is usually to hand waiting to be tested. While unquestionably the age group is slightly more stable and more constant than the grade group, it may still be more practically advantageous for McCall to use an unselected mid-year grade group such as Grade VII than an age group. The burden of proof rests upon those who advocate the use of unselected twelve-year-olds that this gives any added advantage commensurate with the administrative difficulty of securing such a group.

For we cannot too often warn ourselves not to go to too much trouble to standardize. We are always standardizing our test by a criterion which we trust is changing. In other words, we design our tests as part of a larger movement which it is hoped will change the conditions themselves by which we standardize. It is

as though in a cold country where there are no thermometric standards we should attempt with elaborate minuteness to standardize the points on our scale by taking at a particular period of time the average temperatures of a large number of dwellings, simultaneously starting a movement, with the help of the scale thus standardized, to raise the average temperature of the houses. The more effective the movement and the more our scale hastened the change, the more impossible it would be to maintain anything more than a rough standard of temperature. To select our houses too carefully or to spend much time in calibrating our instrument by such a criterion would obviously be absurd. Now this is the exact state of affairs with certain of the scales. Thorndike and McCall recommend their reading scales as devices for improving the objective product; yet it is by this product, a changing factor, that they propose to standardize the scale. For rough estimates no doubt this procedure is justifiable, but it shows how easily too much time may be spent in elaborate standardization. All this is pertinent to the question as to whether an age group which is difficult to procure, furnishes a sufficient amount of increased stability to justify its use.

There is of course no need, when considering the relative merits of age groups and grade groups, to decide exclusively in the favor of one or the other. Both are essential. This paper makes the plea that, all things considered, the grade group is not only more convenient but also more significant in school practice and school thinking. Age groups will, however, have to be used for many purposes and, suggestions are later made for a uniform scheme of presenting results for such groups, similar to that suggested for grade groups. For example, in the matter of intelligence we have every right to think in age groups. Probably the best procedure is to place a child in each age group, and make no attempt to refer him to groups that have had more or less years of environmental experience. Any such comparison is perilous at best. This point is taken up at fuller length in an experimental study on "A Further Criterion for the Selection of Mental Test Elements."[6]

[6] An article to appear in the Journal of Educational Psychology, May, 1922.

It is quite apparent that if the passage is to be made from crude scores to achievement quotients, based on mental age, one must work with age groups. Such passage will, however, if it is to be at all reliable, necessitate very accurate instruments which measure very specifically the two factors involved, namely school achievement and mental achievement. The present attempts in this ambitious direction, involving as they do all the errors characteristic of both types of measurement must be taken *cum grano salis*. Especially is this true when the very tests by which we measure school efficiency are themselves being used with satisfactory results by other experimenters, to measure the other variable—namely, intelligence itself. Thus, for example, the National Intelligence Test employs a fundamentals test as a measure of intelligence. Pintner[1] using a non-linguistic test is less subject to criticism than Monroe and Buckingham; but a careful consideration of the inaccuracy of the two measurements, intelligence and school achievement, together with the fact that we are certainly measuring to no small degree the same factors, should make us very careful in interpreting such results at their face value, for the individual pupil.

The same criticism regarding reliability applies to the work of Monroe and Buckingham in the Illinois Examination. Here crude scores are transmuted into achievement ages, from which they pass to achievement quotients. To overcome the difficulty that the high scoring pupils in the higher grades score more points than the norms for thirteen and fourteen years, the authors are compelled to resort to the hypothetical achievements of age groups which cannot be obtained empirically. This extrapolation is carried almost to the limit when for a particular crude score in arithmetic we find on transmuting it into achievement age that it corresponds to a performance of 25-year-olds.

[1] Pintner says that for 300 children, ranging from the third to the sixth grade a correlation of .47 was obtained between his mental index and the I. Q. on the Stanford-Binet. For the purposes for which he uses his indices, even allowing for the unsatisfactoriness of the Stanford-Binet as a perfect criterion, it seems hardly justifiable, considering the range of the group tested, to say, as he does, that such a relationship is "very satisfactory." If there is the same unreliability for the educational index the ratios must be exceedingly unreliable.

The significance and accuracy of such achievement ages, dependent as they are on assumptions with reference to the normality of distribution or upon extrapolation are much open to question. Seeing that there is so much difficulty in determining the standard scores, for a random group of (let us say) twelve-year-olds, which actually exists and can be subjected to testing, it would seem that any device which demands assumptions with reference to the scores of seventeen and eighteen year groups, could hardly be considered a permanently satisfactory solution. Moreover, there is again the difficulty in interpretation. In considering the achievement in an arithmetic test, how few of us are helped at all by saying that two performances are separated by an amount characteristic of the difference between sixteen year and twenty-three year old standing.

Buckingham[8] has made the suggestion that in the attempt to procure ease of interpretation we should call, by agreement, the median achievement of Grades VIII, VII, VI, etc., 8, 7 and 6 respectively. While this procedure is attractive in its simplicity, it suffers from the great drawback that it is impossible to refer the high scores of Grade VIII and Grade VII to any standard; they exceed 8 and we can only have by doubtful extrapolations a 9, 10, 11, or 12, for hypothetical groups. The suggestion of this present paper is to combine some of the merits found in the devices of McCall, Buckingham and Pintner, particularly the last, with certain other factors to accomplish the end of securing an obvious interpretation of any transmuted score, expressed in terms which make it readily applicable to the school and administrative situation. This suggestion is to transmute the crude scores of the test, into the mid-year distributions of the various grades.

This procedure will be made evident by reference to Table I which was obtained from limited data on the four tests comprising the Chapman Class Room Products Test.[9] In this examination battery the four tests are: (1) arithmetic fundamental

[8] Buckingham, B. R. "Suggestions for procedure following a testing program," *Journal of Educational Research*, 2: 787-801, December, 1920.

[9] Published by the J. B. Lippincott Company.

(A F); (2) arithmetic problems (A P); (3) reading selections (R S); (4) reading continuous (R C). For each of these tests and for the total for Grade V, VI, VII, VIII the achievement at mid-year for the nine equally increasing (with generous assumptions!) levels, 0.4 sigma apart, are shown. These levels are given the descriptive designation shown in the first column, and the quantitative designation ranging from 1 to 9, shown in the second column. Whether mid-year or end-year standards are chosen is a matter of agreement. For tests of skill which are the result of long training, the former is advantageous; for tests of informational material which is acquired in a single grade the end-year standard is referable. Assuming mid-year standards as given, suppose a pupil in Grade VI scores on the total test 47 points. By referring to Table I, it is possible to determine the levels of achievement for Grade VI, VII or VIII. In the case considered, it will be seen that the particular pupil's achievement is sixth-grade high or VI-7. This will be the usual status wanted. Where, however, for purposes of regrading the desire is to discover the place of the pupil in Grade VII or Grade VIII, it can be readily seen from the table that his score corresponds to seventh-grade median or VII-5, and to eighth-grade medium low or VIII-4. This illustration brings out an important merit of the proposed system of reporting. Those who are working in the schools with these measuring instruments are trying to break down certain ideas concerning the present system of grading, particularly the notion that a grade represents a narrow level of achievement of a homogeneous group. The most powereful instrument to employ in attacking this problem is one which shows on its surface the extent to which overlapping of the grades is present. This the suggested system is well calculated to accomplish; for when the transition is made from the crude scores to the equivalent scores, a momentary examination of the table shows that the same crude score may be characteristic of different levels of achievement in different grades. The proposed scheme focuses attention upon grades and the wide range of achievement in a grade, rather than on a distribution in terms of

TABLE I.—NINE LEVELS OF ACHIEVEMENT FOR EACH GRADE AT MID-YEAR. CHAPMAN CLASSROOM PRODUCTS SURVEY TESTS.[a]

Levels of Achievement			Grade V					Grade VI					Grade VII					Grade VIII				
Descriptive Designation	Quantative Designation	Percentile (approx)	A. F.	A. P.	R. S.	R. C.	Total	A. F.	A. P.	R. S.	R. C.	Total	A. F.	A. P.	R. S.	R. C.	Total	A. F.	A. P.	R. S.	R. C.	Total
Indeterminate	9+		13+	10+	12+	13+	38+	17+	15+	15+	18+	57+	18	21+	16+	20+	67+	18	22+	18	22+	71+
Highest	9	95	12	9	11	12	37	16	14	14	17	56	17	20	15	19	66	18	21	17	21	71
Very high	8	88	11.5	8	10	11	34	15	12	12	16	51	16.5	17	14	17	61	17	18	16	20	65
High	7	79	11	7	9	9	31	14	10	11	14	47	16	15	13	16	56	16.5	16	15	18	61
Medium high	6	66	10	6	8	7	28	13	8	10	12	42	15	13	12	14	52	16	14	13	17	57
Median	5	50	9	5	6	6	25	12	7	8	10	38	14	11	10	12	47	15	13	11	15	52
Medium low	4	34	8	4	4	4	22	11	6	6	9	34	13	9	8	10	43	14	11	9	13	47
Low	3	21	7	3	3	3	19	10	5	4	7	30	12	7	6	9	39	13	9	7	11	43
Very low	2	12	6	2	2	2	16	9	4	3	5	26	11	6	5	8	35	12	8	6	9	38
Lowest	1	5	5	1	1	1	13	8	3	2	3	22	10	5	4	5	31	11	6	5	8	34
Indeterminate	1–		4–	0	0	0	12–	7–	2–	1–	2–	21–	9–	4–	3–	4–	30–	10–	5–	4–	7–	33–

[a] A. F. = arithmetic fundamental; A. P. = arithmetic problems; R. S. = reading selections; and R. C. = reading continuous. A pupil of Grade VI scoring 10 on reading selections and receiving 47 as a total score has his scores recorded as VI-6 and VI-7 respectively. If reference is to Grade VII, these scores are both recorded as VII-5.

a single population. Of course, the wide variations of achievement on the continuous scale show the same thing, but do they show it as obviously? We have already made the point that the thinking of the practical schoolman is in terms of grades, and that eventually children, as the result of the tests, have to be placed in grades. There is, however, a further reason why standardization by grade groups rather than by age groups should be given preference. This is found in the fact that the subjects in which we test are taught by grade rather than by age. By this is meant that a certain part of history or civics is taught to Grade VII and Grade VIII and not to 13- or 14-year-old groups. It seems disadvantageous to use as the standardizing group a population such as McCall's of which a considerable portion have never been exposed to the instruction in the elements out of which the test is constructed. While this criticism applies much less to examinations in reading, and to wide examinations in arithmetic, yet in specific elements in arithmetic taught in later grades it is certainly very pertinent.

We shall make no attempt to pass above the 95 percent point which is called 9, because the data, unless originally very great in amount, are becoming so attenuated at this point that reliability is out of the question.[10] Where the score passes beyond that of the 5th or 95th pupil at the extremes, it may well be left indeterminate, calling it 1— or 9 +. To refuse to pass beyond these points and to leave the rating indeterminate if less satisfying is at least an intellectually honest procedure.

For each grade there are thus nine levels each separated by equal amounts. These for any particular grade, such as VI, are written VI-1, VI-5, VI-9, etc. Together with the indeterminate levels, VI-1 — and VI-9 +, there are thus eleven levels for each grade. If the data are sufficiently homogeneous and extensive to justify the process, we may insert intermediate steps with suitable descriptive terms and with quantitative designation, e. g., 1.5, 2.5, 3.5, etc., up to 8.5. This would furnish nineteen levels of achievement.

[10] Yule, G. Udney. *An introduction to the theory of statistics.* (Fourth edition.) London: Charles Griffin and Company, p. 341.

When very fine correlation work has to be done with the transmuted scores, it will of course be necessary to work to smaller limits than those afforded by the eleven or even the nineteen levels. This however will be a simple matter of interpolation and extrapolation. We can readily decide not only the achievement of VI-5 or VI-6, but also the achievement of VI-5.2, VI-5.3 or VI-5.4.

When dealing with an extreme case falling beyond the equivalent of 9 for a particular grade, it is surely better to pass to a higher group and place the case in that, than to rely on very inaccurate points at the end of a distribution, as McCall is compelled to do. If a pupil in Grade VI falls in VI-9 +, it is obvious that he is out of place, and that by referring to the distribution of the higher grades, we can see where he fits. If, however, we are giving tests to measure improvement, the scores of the few pupils at the two extremes would have to be given special treatment.

When it is desired to work in age groups, the same procedure as that above advocated would be recommended. Here the prefix would be in Arabic instead of Roman numerals, e. g., 12-5 would indicate the median achievement of unselected pupils in the neighborhood of 12 years old. Roman figures would always indicate grades, and Arabic figures would indicate ages. Meanwhile, seeing that almost all thinking takes place in grades, particularly when the problem is that of regrading, it is well to establish the grade norms first rather than the age norms. If, however, we wish to make comparisons from country to country, there is much to be said for age norms. Certain English writers, for example Ballard, are making the appeal for age norms for the simple reason that they do not understand the American grading systems.

In the matter of intelligence tests where primarily the thinking is done in age groups, the above convention would also apply. The Arabic figures would represent the age and the figure following would indicate the level of achievement. In certain cases where the measures of intelligence were very reliable, it might

be necessary to establish by interpolation or experimentally intermediate achievements. For example for pupils 10 years, 6 months the designation might be 10.5-7.

The same convention as to grades could be extended beyond the elementary school, the various years of high school being given the designation IX, X, XI, and XII, and the corresponding levels for each high-school subject determined.

When the main interest is in measuring improvement, particularly over short spaces of time, rather than in regrading, the equally spaced steps of any of the various grade levels may be used. Other things being equal, grade VII levels would probably give the most useful system of units. In those cases where the desire is to measure over a longer period, say a number of years, two alternatives present themselves. In the first place, we may measure the improvement discontinuously in terms of the different grade levels. An advancement from VI-5, to VII-5 would represent a year's improvement, in the same way as an advance from VII-7 to VIII-7 would represent, neglecting selection, a normal year's growth. In fact, when measured in these terms ease of interpretation and usefulness of information are at a maximum; for when merely grade averages are reported, as a test is repeated from year to year, little attention is given to those that exceed the average, largely because there is no ready method of checking up their degree of improvement. When, however, we have the various levels for the different grades established, interpretations of degrees of improvement of pupils far removed from the median will be readily possible. Thus a score increase from VI-9 to VII-9, will be just as good for a pupil of the nine level, as a jump from VI-5, to VII-5 will be for a pupil of the median level.

If, however, a continuous scale is demanded—and physical measurement suggests that in those few tests which are applicable to a wide range of grades, such a scale is the ideal—I would suggest that we use the total population of grades five to eight, employing unselected children at mid-year from these four grades in equal number. Considering these groups as a single population,

we may determine the nine levels, or for finer grading the nineteen levels, in the same way as with a single grade group. It would be necessary to have some designation for such a series of steps. Perhaps U. S. might be used, the letters standing for "Universal Scale."

Such a combination of 100 unselected pupils from each of the four top grades at mid-year has been made for the above mentioned test, with results as shown in Table II. This combination of 400 is treated as a single group, and the nineteen levels of achievement in the totals of the four tests are exhibited. In this way we get a scale which is continuous in the same way as McCall's but which is less subject to the censure that extreme scores cannot be scaled. Furthermore, it is derived from a population which can readily be obtained for testing purposes. Only in a few of the wider tests, reading *par excellence*, could this broad population be used. Otherwise the procedure would be subject to the same criticism as is McCall's, in that we should be standardizing our scale on a population a large section of which had been given very little or even no chance to score in the test. The fact that in the narrower tests we are compelled to use short scales, which, so to speak, break at various points, is probably inherent in our problem. The physical analogy of the yard stick, while suggestive, is not really applicable. There is continuity, for example, in height within any group, say twelve-year-olds; but the same is not true for knowledge of civics, or of interest in arithmetic. Here we strike the problem of the undistributed zero.

To avoid misunderstanding, it is perhaps well in closing to point out explicitly two demerits of the universal scale procedure. The first is found in the experimental fact that grade groups are less constant than age groups in achievement. This is caused by differing procedures with reference to promotion from system to system. The second disadvantage is that in dealing with grade groups, standardization must take place at some definite point in the history of the grade, presumably either at mid-year, or at end-year. In contrast to this limitation, standardization by age

groups can in the general case be done at any time. Neither of these drawbacks is sufficient to prevent the procedure from being used. The first point has already been met; whatever norms are obtained will always be subject to greater inaccuracy than will be produced by varying standards of promotion. Probably care

TABLE II.—A UNIVERSAL SCALE[a]

Descriptive Designation	Per-centile (approx.)	Quantitative Designation, Equivalent Score	Crude Score
Indeterminate highest	..	US-9+	68+
Highest	95	US-9	67
	92	US-8.5	63
Very high	88	US-8	60
	84	US-7.5	57
High	79	US-7	53
	73	US-6.5	50
Medium high	66	US-6	46
	58	US-5.5	43
MEDIAN	50	US-5	40
	42	US-4.5	37
Medium low	34	US-4	34
	27	US-3.5	31
Low	21	US-3	28
	16	US-2.5	25
Very low	12	US-2	23
	8	US-1.5	21
Lowest	5	US-1	18
Indeterminate lowest		US-1-	17-

[a] Totals in the Chapman Classroom Products Tests at the various levels. Population, 100 unselected pupils from each of Grades V to VIII. A pupil scoring 53 is recorded as US-7.

will have to be taken to secure standardization material from school systems in which the age-grade tables are reasonably normal. The second point raises a real difficulty but not as great as that which will have to be met in combining age groups, a procedure to which I believe, McCall will have to resort if his scale is to be at all reliable at the extremes. To secure unselected twelve-year-olds is within the bounds of possibility, but the higher unselected age groups can probably not be attained and the low unselected age groups will bring in the problem of the undistributed zero.

Summary

This paper presents a scheme of reporting test scores whereby, after the transition has been made from the crude score to the equivalent score, the latter bears, upon its face, an obvious and useful interpretation.

Nine equally separated levels of achievement for each grade are set up. These levels for a particular grade being referred to as VI-4, VII-8, etc. In addition to this quantitative rating, these nine levels are given a descriptive designation.

For age groups, the same method of report applies, Arabic numerals replacing the Roman numerals.

Intelligence levels at all ages admit of treatment by the same method.

Where a continuous scale is demanded, the mid-year Grade-VII scale is suggested. If this is not wide enough, an unselected population consisting of the four top grades (at mid-year) in equal number is recommended. This is treated as a single group and the same nine levels determined. If the data are extensive, these nine may be extended to seventeen or much further. The designation for a level of such a group is US-5, US-8, etc.

In every case the figure or letter before the hyphen signifies the group to which reference is made, the figure after the hyphen indicates the level attained.

AN ACCURATE INDEX OF NATIONALITY

RIVERDA HARDING JORDAN
Cornell University

Many studies published during the last decade have included in their findings data relative to nationality, and from these data some wide-spread conclusions have been drawn. In most cases, the authors of these studies have not themselves been disposed to attach too much importance to their findings, since they realize the imperfect character of the data; but the serious factor has been that these findings have been quoted aside from the context, and so have given rise to inferences not always justified by the facts. Therefore, when it becomes possible to point out methods of procedure productive of more accurate results, a check is obtained by which these inferences from past studies may be verified or invalidated.

Certainly, one of the most important elements entering into the whole matter of nationality investigation is the reliability of the index used to establish the nationality factor. In school surveys, the usual method has been to determine the nationality of the father of the school child; and the assumption has been that the nationality of the child is exactly determined by the birthplace of the father. Now several questions immediately arise as to the validity of this method, especially when, as in the case of school children, it is desired to determine the general influence of the home background. Two of these questions seem especially pertinent: first, does the birthplace of the father really indicate the national background of the child; and second, does the father's influence in the home determine the language spoken in the home, so that it is fair to conclude that the language there spoken is the native tongue of his place of origin?

In working out a recent study of the relationship borne by various nationalities of school children to their progress,[1] data were collected which make it possible to give a partial answer to

[1] R. H. Jordan, *Nationality and School Progress*. Public School Publishing Co., 1921.

these questions. The investigation was carried on in the cities of Minneapolis and St. Paul, where the population is sufficiently polyglot to justify considering these cities as typical of the inland centers of population. In all, 2653 cases of school children were studied, and data were collected with reference to the birthplaces not only of both parents of each child, but also of the four grandparents. In 2607 cases the birthplace of the father was determined, and in 2491 of these, of the four grandparents as well. The nationality of the child was based upon the place of origin of his grandparents, according to the following method. The birthplaces of the parents were disregarded, and the child was listed as belonging to that nationality shown by the predominating number of agreements of his grandparents. That is, if all four grandparents were born in some one foreign country, the child was listed as being of that nationality; or if three were of one nation and one of another, or two of one nation and the other two of different nations, the majority of grandparents determined the nationality of the child. But where there was no predominance, as where two were born in one country and the other two in another, the child was regarded as of "mixed" nationality, and so was not included in tables of comparison. To make the matter clearer, if all four grandparents were Norwegian, or if three were Norwegian and one Swedish, or if two were Norwegian, one Swedish, and one American, the child was listed as Norwegian. But if two were Norwegian and two were Swedish, he was listed as of mixed origin. Thus, a comparison can be made between the results secured by listing children according to the birthplaces of the fathers, and the results obtained from using the places of origin of their grandparents. Table I shows the comparative results.

No close analysis of this table is necessary to make clear two points: first, that there is a fairly close agreement between the foreign-born groups, and second, that there is no agreement at all between the American-born groups. Accordingly, it may be inferred that the birthplace of the father is a fairly accurate index of the origin of the child if the father is foreign-born, but

TABLE I.—NATIONALITY OF CHILDREN ACCORDING TO BIRTHPLACE OF FATHER AND ACCORDING TO BIRTHPLACE OF TWO OR MORE GRANDPARENTS

Birthplace	Fathers	Two or More Grandparents
United States	977	350
Great Britain	149	165
Germany	104	177
Norway	209	215
Sweden	422	420
Austro-Hungary	142	129
Roumania (Jew)	115	99
Russia (Jew)	340	336
Poland	22	38
Finland	43	45
Denmark	34	22
Italy	19	18
Miscellaneous	31	20
Mixed origin, no predominants	..	457
	2607	2491

that it is not at all an index if the father is born in the United States. Therefore, whenever there is any comparison of foreign nationality with Americans, the father's birthplace cannot be taken as a fair index of the real ancestry of the child; and such comparisons must be based upon more accurate information. But when comparisons are made between national groups where Americans are excluded, the birthplace of the father may be considered as a rough index of the child's ancestry. An objection is frequently made that it takes too much time and trouble to obtain the data relative to grandparents; but it may be pointed out that the method ordinarily employed, namely, having a card filled out at home by the child's parents, will involve but little more time or trouble, if the parents are asked to include as well the birthplaces of grandparents.

A study also was made of the relative influence of the father and the mother in determining the language spoken in the home. The cases were classified into five groups as follows:

1. American born father, foreign born mother:
   - Both speak English at home 84
   - Both speak foreign language 13
   - Father speaks English, mother her native language 14

   —— 111

2. Foreign born father, American born mother:
   - Both speak English..... 139
   - Both speak foreign language..... 61
   - Father speaks native language, mother speaks English 9

   —— 209

3. Both foreign born, one of one country, one of another:
   - Both speak English..... 52
   - Both speak father's language..... 21
   - Both speak mother's language..... 13
   - Father English, mother native..... 5
   - Father native, mother English..... 5
   - Each original language, unchanged..... 13

   —— 109

4. Both foreign born, in the same country:
   - Mother speaks English, father native..... 6

5. Both born in the United States:
   - Father speaks English, mother a foreign language.... 2
   - Mother speaks English, father a foreign language.... 1
   - Both speak a foreign language..... 21

   —— 24

Total Cases 459

A summary of the above cases shows that in 166 the father has seemed to dominate since his native language has been adopted as the home language; and in 165 the mother has seemed to have the greater influence. In 36 cases the original differences of birth have persisted, there being "no compromise." In 55 cases the father and mother speak different languages in the home. Of the 115 foreign-born cases where both parents are foreign-born, only five fathers speak English at home when the mothers retain the native tongue, but eleven mothers speak English while the fathers use their native language.

Naturally, the conclusions to be reached are largely colored by two great factors, first, the natural tendency to pick up and to use English in an English-speaking community, and second, the tendency to continue the use of the foreign language where grandparents are resident in the home. In the latter case, the influence of the father might frequently be neutralized by the presence of his wife's foreign-born parents, and vice versa. However, these factors do not affect the main point, *i.e.*, that the

evidence shows rather conclusively that the father is by no means to be assumed to be the predominating influence in determining the atmosphere of the home. Rather must it be concluded that the mother is at least quite as influential, with the general preponderance in her favor. All of this seems to demonstrate that there is no good ground for assuming that the birthplace of the father is any clear indication of the language spoken at home.

From these two studies, there seems to be no doubt that for the groups studied, national comparisons drawn on the basis of the father's place of birth would be very largely erroneous and misleading. Inferentially, a similar doubt would be placed upon all conclusions drawn from data collected without further knowledge than the origin of the father. Since but little more time or trouble will be involved in getting all of the needed data, is it too much to ask that in future, the studies of school children involving nationality be based upon the birthplaces of their grandparents?

Editorials

WHAT IS A PUBLIC SCHOOL

In England nine institutions have been known for many years as the Great English Public Schools. They charge exorbitant fees and cater to the *élite*. Satirically—and with the half truth that makes satire effective—it has been said that they are called *great* because they limit their ministrations to a few highly-selected students, *English* because Greek and Latin constitute the core of instruction, *public* because none but the rich can attend, and *schools* because the chief interest is football and cricket. A few institutions in this country resemble these aristocratic corporations. But we do not call them public schools; and not many of us think of them as great. English, however, they may be.

In this country a public school is free and tax-supported. But it should be much more. It should serve the public and should be understood and admired by the public.

Not long ago the writer visited a school system in a peculiar community. A fine progressive spirit was evident among the teachers and supervisors. Modern buildings were either available or under construction. Good teaching was being provided. The schools, however, were isolated. Their lines of communication with the public were insufficiently maintained. No parent-teacher organization was in evidence. Though the teachers were carefully selected and manifestly of good ability, their contacts with the community appeared to be few. They seemed to be living unto themselves. Their social talents, which were of a high order, were exhibited for the most part to each other.

Such a school falls short of being "public" in the best sense. It is likely either now or later to suffer from a lack of the material support without which good teaching, good buildings, and good

equipment cannot be purchased. Moreover, it will have little of that larger influence upon the life of the community which is at once its privilege and its duty. Children will be sent to school because they must be sent; but, they will be withdrawn as soon as the compulsory age is passed, and during their attendance the teachers will be unable to count on parental support. With the school and home working independently of each other or at cross purposes, even the narrower objectives of the school will be relatively unattainable. The progress of pupils will be slow, and elimination will be high. Boys and girls too old for their grades will be numerous, and the membership of the high school will be small.

On the other hand, the public will be deprived of the good influences of a real public school—of a school which reaches out into the community, shares its best endeavors, and gives momentum to its best impulses. In the district where the school is conducted behind closed doors, the business man complains most loudly that children can no longer spell or add or write grammatically. The "best citizens" and their wives regard school teachers as socially impossible, while by others they are regarded as snobs. And the pity of it is that both characterizations tend to become true as the school continues to hold itself aloof from the people.

The public, and the school must, therefore, be brought together. The public is really tremendously interested—or will be if its attention is called to its interests. It provides both the money and the children; and these constitute at once the essentials of the school and the people's dearest possessions. It would seem, therefore, that only through the existence of really unnatural conditions can the school and the people be estranged or even become indifferent to each other. Wherever such conditions have insensibly developed, it should be the first care of school people to banish them.

That school is truly a public school which is not only free and tax-supported but which is also operated as a high public trust. Such a school will render continuous account of itself to the pub-

lic; it will take the public fully into its confidence; it will assert and maintain, on the basis of undisputable worth, its right to lead and its ability to inspire. When a crisis comes—when agitators or obstructionists arise, or when business depression reduces valuations—such a school will be protected. Aristocrats will send their children to it instead of to exclusive private schools. Fathers and mothers—aristocratic or otherwise—will wish their children to enter it at an early age and remain in it as long as possible. They will speak well of the school and second the efforts of teachers. Undoubtedly one of the best ways to reduce retardation and elimination is to begin with the parents. Many a school system which has been surveyed and found wanting needs nothing so much as to become a real public school.

B. R. B.

THE PRINCIPLE OF RANDOM SAMPLING IN SUPERVISION

One big difficulty in supervision is that the supervisor is not ubiquitous. He cannot be here and there, too; and if he is here today he must be there tomorrow. If he has lately given attention to arithmetic in the sixth grade, the next time he comes around he will have to hear some reading or look over a set of spelling papers. It is hard to give consecutive attention to any particular type of work long enough to effect improvement.

Standardized tests, however, have multiplied the hands of the supervisor. He may not appear in the classroom as often as he used to—nor as often as he would like to; but he may be keeping track of things far better than formerly. Contemptuous reference to the "office superintendent" means nothing unless we know what he is doing in his office. He may be making effective use of some of the devices of statistics, either in conjunction with standardized tests or without them. We called attention in a recent issue to an application which he might make of the median.

Another statistical principle that will help the supervisor is that of random sampling. Let him rate the forty pupils of a class in any subject, and then let him select at random the ratings

of ten pupils. He will probably be surprised to find how closely the average of the ten ratings corresponds to the average for the entire class. He may therefore use the ratings of the ten as a fair substitute for the ratings of the forty unless individual ratings are required. It is true that a random selection occasionally yields a result which differs materially from the true result. But the ease with which the method may be applied makes it practicable to obtain many determinations. Since the vast majority of these will be close to the true determination, no serious error will be involved.

Suppose a principal is seeking the most effective way to stimulate pupils to greater skill in written expression. Undoubtedly, he will entertain the idea that if the effort could be kept up, the maximum result would be obtained by securing a little writing every day from every pupil. But he will readily see that for a teacher to read and to rate a set of compositions every day, even if they were very short, would be too much to expect. At least it would be too much to expect unless he were willing to allow other subjects to be neglected. And even if a teacher could rate a short composition a day from every child, it would be clear to the principal that it would be impossible for him to review more than a very few of these papers, perhaps not more than those of three or four classes daily.

By the use of the method of random sampling, however, and with a little planning, the principal may attain surprising results in the personal supervision of English composition. The following method has been tried. Let from ten to fifteen minutes per day be devoted to English composition by each class. Such a part of that time as is not used for discussing previous work should be devoted to the writing of two or three sentences or a paragraph on an assigned topic. After collecting these exercises, the teacher will simply endorse them with the name of the class, the date, the number of papers, and the topic on which the children have written. A pupil messenger will call for the papers together with those of other classes in the building and will take them to the principal's office. Here the same pupil, or another one, will

enter in a record book the facts which the teachers have endorsed on each package. This pupil will also draw from each package a certain number of papers (say ten). These papers are the random samplings from an inspection of which the principal will obtain a fairly accurate estimate of the work of each class for a given day. It is suggested that he read each of these ten papers for each class and that he rate each paper as satisfactory or unsatisfactory. It will in general be found unnecessary to recognize more than these two classifications—namely, satisfactory or unsatisfactory. If seven of the papers are satisfactory, the class rating is seventy. If eight of them are satisfactory, it is eighty. After the random samplings for each class have been examined and rated in the manner indicated, the principal will indorse the rating on each package and enter it after the name of the class in the record book. The packages will then be returned to the teachers.

By this means, it has been found possible for a principal to give close attention to the work of as many as twenty-five classes without an undue expenditure of time. The amount of improvement in English writing which has resulted from the use of this device has been exceedingly gratifying.

This, however, is only an illustration. What we wish to convey to the reader is first, that a number of statistical principles may be used by the teacher and especially by the supervisor with very gratifying results, and second, that the method of random sampling is a device which is particularly effective and distinctly economical of time.

B. R. B.

Reviews and Abstracts

JORDAN, ARTHUR M. *Children's interests in reading.* (Teachers College, Contributions to Education No. 107) New York: Teachers College, Columbia University, 1921. 143 pp.

This is a study furnishing facts of deep significance to all who are in any way concerned with the reading of children from nine to eighteen years of age. In stating the purpose of the study Dr. Jordan writes as follows:

"The abiding interest of school children in their subjects of study has for many years been one of the criteria of good instruction. In classes in literature this becomes a problem of peculiar interest. . . . If we could determine what the child's major interests are, be those interests good or bad, it would be possible to direct these forces along lines which are desirable. If not directly, at least indirectly, we could connect the subject with his interests and show how this subject is related to these interests. Psychologically this interest would strengthen the connections between the bonds formed in the reading of good books. Thus what has been learned would be longer remembered. In the second place there would be more possibilities for secondary neurone connections, that is, interest in reading would lead on to greater interest."

After a review, comparison, and criticism of previous studies along this line, the author explains his investigations by the questionnaire and library methods. The reader is impressed throughout with the care with which the study was planned and with the scientific methods used in carrying it out, in compiling the data, and in arriving at the conclusions.

Most striking evidence is given of the pronounced sex differences in the choices of books and magazines. The tables show also the strength of the various interests from year to year and the relative popularity of the most popular books as influenced by sex and by age. In a careful interpretation an analysis is made of the causes of these differences. The correlations obtained between the results of the different investigations show clearly the large differences between the sexes in their liking for books—these differences being greatest at the ages of twelve to thirteen in the cases of both books and magazines. "The magnitude of these differences is, finally, in no small degree determined . . . first by the boys' great interest in books of adventure and in the magazines of adventure and of science; and by the girls' preference for books of fiction and for magazines of fiction and woman's arts."

Altogether, this is an illuminating study of which extensive application should be made in literature courses, in the compiling of reading lists, and in the writing of books for young people.

MYRTLE L. KAUFMANN

Elementary Supervisor, Logansport, Indiana

TRYON, ROLLA MILTON. *The teaching of history in the junior and senior high schools.* New York: Ginn & Company, 1921. 294 pp.

This book on the teaching of history in the junior and senior high schools fulfills a long felt want for a book bringing together many concrete suggestions and devices for the teacher of history. It will be particularly helpful for the beginning teacher who faces the problem of organizing several courses in the various fields of history.

The chapters giving suggestions on methodology are very helpful. They describe at some length the history recitation; detail ways and means of teaching pupils how to study history; analyze and evaluate special methods of procedure such as the lecture, topical, textbook, source, and problem methods of presenting subject matter. The book also comments upon written work in history, such as, the permanent notebook, the collateral reading done in the library, and the term paper.

Some of the other chapters in the book are not as concrete as those dealing specifically with methods of teaching. For example, the section on Progress within the Subject, does not clearly set forth differentiating principles for the allotment of subject matter to each grade. The chapter on Measuring the Results of History Teaching would have been of more service if it had suggested which of the standardized history tests to use, and how to interpret results so obtained.

A very important section in the book, particularly for the new teacher, is the one on Planning the Course and the Lesson. In the opinion of the reviewer this chapter could well have been expanded. Concrete proposals are made concerning the main and subordinate topics to be covered in the course and illustrations of daily lesson plans are given. They indicate the importance of a daily outline including important events, persons, and dates; further references; maps; and pivotol questions and exercises. Such suggestions should stimulate teachers to work out an efficient method of organization and also make the daily teaching of history more pleasant.

It is to be regretted that certain of the much discussed psychological aspects of history teaching are not discussed in the book. We need a discussion concerning the place of association in history teaching, and answers to such learning problems as (1) how much repetition is necessary to "fix" important historical facts? (2) How can we stimulate and train the imagination and judgment in the teaching of history? The little research done thus far on the psychology of history has not been reported in sources available to the average teacher. Books presenting material on the teaching of history should at least summarize this data.

The reviewer agrees with the statement in the preface of the book that while there are many ways of teaching history, it is fundamental to educational method that the teacher know that there are a number of ways of doing a multitude of things connected with the everyday procedure of that subject. This book tells the teacher of history about many of these ways and points out a procedure in managing them.

EARL W. RUGG

Columbia University

DAVIS, SHELDON EMMOR. *The technique of teaching.* New York: The Macmillan Company, 1922. viii + 346 pp.

This book, like several of its type, has for its purpose the improvement of teaching. The preface states: "Its professional purpose is to analyze classroom activities which the instructor may have performed mechanically without being conscious of principles involved, to afford opportunity for intelligent evaluation of procedures in common use, and to suggest as diversified methods of presentation as the limits of one small volume render possible."

The first chapter discusses the general principles of the technique of teaching. Each of the remaining six chapters is devoted to the teaching of one of the fundamental subjects, as follows: spelling, reading and literature, composition and grammar, arithmetic, history, geography. The first chapter gives brief discussions of several well-known principles. It is doubtful whether this chapter will be of much value to the novice owing to the brief and sketchy manner in which the principles are discussed. The well-trained teacher is already familiar with such conceptions and the discussions will add little to her stock of ideas. It seems to the writer that the book would be of more value if it contained a more elaborate discussion of the principles in light of classroom activities. A teacher who has been performing these activities mechanically would scarcely gain sufficient background from a study of this book to interpret the principles underlying such instruction.

It is surprising to find so little attention given to recent experiments in education. The discussions on teaching the various subjects are quite similar to those in several books that have been in rather wide use for a number of years. The justification for a new book on methodology would be that it would incorporate the results of our latest and most comprehensive experiments. Owing to the scope of the book, the discussion of the teaching of each subject is much less detailed than one finds in books which are entirely devoted to the teaching of one specific subject. This, however, could be justified if the material presented were of a different type. The exercises at the end of each chapter are deserving of special commendation. They show evidence of careful thinking. They should be provocative of thought and would furnish an excellent basis for class discussion. From the writer's point of view these exercises are the most valuable portion of the book. If the book succeeds in interesting more people in the improvement of the technique of teaching, its publication may be justified. Aside from the exercises, however, it is questionable whether the book is a contribution to the teaching process.

C. R. MAXWELL

University of Wyoming

PRYOR, HUGH CLARK and PITTMAN, MARVIN SUMMERS. *A guide to the teaching of spelling.* New York: The Macmillan Company, 1921. 151 pp.

The authors have prepared a manual on the teaching of spelling primarily for the inexperienced teacher. It consists of three main divisions: fundamental principles; methods, materials, and devices; and a bibliography.

From the viewpoint of the major function of the manual, the most valuable features are the chapters on the "best methods," on devices, and on ques-

tions often asked by teachers. These features are discussed in simple language in terms of the underlying principles that have been recently formulated by scientific researches. This method of presentation is commendable because it tends to keep the amateur teacher from the rut of "blind imitation."

The chief negative criticism of the three foregoing desirable features of the manual is the meagerness of illustration. For example, it is questionable whether there are enough illustrations of the "best methods" to make each one of them clear to a beginner in the field. Not only from the standpoint of clearness, but also from that of the felt needs of the amateur teacher it would seem that more illustrations might well have been given of the "best methods," of the materials, and of the devices. The popularity of the publications devoted to devices supports the latter reason for more illustrations.

There is a tendency for the authors to cater to the technical demands of the investigator in the field of spelling. For example, note the detailed presentation of "A Suggested Minimal Spelling List" on pp. 79-85 and the extended bibliography of 78 titles at the end of the manual. Such lapses from the primary purposes of the manual would not seem to enhance its value to the inexperienced teacher.

May we offer two other negative criticisms? First, we fail to see why the chapters were not numbered consecutively. Second, minor errors or editing have been overlooked. For example, in the Table of Contents, p. viii, the Bibliography is not given as Part III, where as it is so given on p. 137; and on p. 28 the word "Interest" in black face should be an "inset" in the last paragraph instead of in the one in which it appears.

However, regardless of these criticisms the inexperienced teacher will be well repaid in suggestions gained by reading the manual.

A. G. CAPPS

University of Missouri

THORNDIKE, EDWARD LEE, *The new method in arithmetic.* Chicago: Rand McNally and Company. 1921. 260 pp.

A long time ago we all learned from our readers that for want of a nail a shoe was lost; for want of a shoe a horse was lost; etc. In this book Professor Thorndike is not too big to be above the nails of arithmetic. Nor does this by any means imply that ample recognition is not given to thinking and applied superstructure. Some given to dialetics may wish to argue whether or not the school should be more cultural or more disciplinary, more academic or more vocational, et cetera, but the implications of the foregoing have validity only if and when children can add, subtract, multiply, divide, measure and weigh, handle commonly used fractions, compute interest and taxes, discriminate in investments, and intelligently appreciate numerical relations in the problems of farm, industry, business, and travel.

"The New Methods" exemplified are based upon "principles discovered by the psychology of learning, by experimental education, and by the observation of successful school practice." The chief concern of the author is not a defense of past procedure but a vigorous advocacy of the best principles derived from careful experimentation. We all know how comfortable the old shoe is,

but there comes a time when we must step out with the new pair. The Thorndike Arithmetics (Rand McNally) are used throughout for illustrative purposes but "transfer" has a mighty effect with any modern set of good texts. The book is constructed from the point of view of the working teacher who may have little direct knowledge of psychology. There is abundance of detailed illustration and application and further end-of-chapter exercises. College and Normal classes should find much dynamic material in "New Methods."

Chapter headings give intimation of the pragmatic character of the book and the following are typical: Reality; Interest; Theory and Explanations; Habit Formation and Drill; Solving Problems; Teaching as Guidance; Some Hard Things; Some Common Mistakes; Some Instructive Disputes; Tests and Examinations. Under what simpler or more sensible doctrine could a method book in arithmetic be launched than with the prelude: "The older methods taught arithmetic for arithmetic's sake, regardless of the needs of life. The newer methods emphasize the processes which life will require and the problems which life will offer."

"Reality" is sensed in: "The newer methods set a higher standard in the selection and construction of problems, requiring not only that they give the pupil an opportunity to think and to apply arithmetical knowledge, but also that they teach him to think and to apply arithmetic to situations such as life may offer, in useful and reasonable ways, and so to esteem arithmetic not only as a good game for the mind but also as a substantial helper in life's work." In fact the problems which "could occur in real life only in an insane asylum" are conspicuously absent from Professor Thorndike's treatment. Note the very significant contrast however when 2 times 2, 3 times 2, etc., and $4 \div 2$, $6 \div 2$ are taught as tables in connection with pints and quarts and the cost of postage stamps. And goodbye for "Denominate Numbers" as such in a separate suite. Old Messrs. Foot, Yard, Pint, Quart, and Gallon meet, mingle and strike up lasting acquaintance "bonds" right in the lobby. But when prohibition becomes firmly fixed, Professor Thorndike will be the very first to snub, arithmetically speaking, the more out-of-date members of the above family.

Two potent interests are recognized: interest in mental activity as such and interest in achievement. While the newer methods increase the strength of these two appeals, arithmetical thinking should be preserved from waste and failure due to difficulties of vocabulary and construction. Interest through achievement is most laudable, with definite tasks definitely set and with results as definitely measured and noted.

The older methods made much use of explanations based upon deductive reasoning and beyond the child's grasp. The newer methods attempt to secure real understanding of rules and principles but by an *inductive* reasoning based upon the simple premise "because by so doing I get the right answer," such answer however being submitted to verification. "The newer methods lay more stress on the pupil's surety that the rule or process is right and less stress on his ability to state in words a proof that would satisfy a mathematician." The child verifies the results obtained in $6 \div \frac{3}{4}$ (after multiplying 6 by $\frac{4}{4}$) by dividing a 6 inch strip into $\frac{3}{4}$ inch lengths. The rule that he formulates is

verified by further examples. "Such a child understands in a certain true and useful way the reasons for 'invert and multiply' or 'multiply by the reciprocal' (supposing him to have been taught the meaning of reciprocal). He may not be able to state the deductive proof from the nature of a fraction. Neither can some of my readers perhaps!"

And so throughout the book one finds many "satisfying connections" in the Law of Effect. There is no saner or more scientifically arrived at arithmetical procedure than the progressive teacher will find in New Methods. The "examination" which really counts is fifty years long and its situations are as real and varied as are the conditions of life. Such an examination however "demands a mastery of a few real things rather than a 60 percent efficiency with many." These real arithmetical processes have no better presentation than in Thorndike's New Methods. No up-to-date school man will long be without a copy.

CYRUS D. MEAD

University of California

HOPKINS, L. THOMAS, *The marking system of the college examination board.* Cambridge 38, Mass.: The Graduate School of Education, Harvard University. 1921. 15 pp.

In a monograph of commendable brevity Mr. Hopkins sets forth the results of a study of the marks given by the College Entrance Examination Board during the period 1902 to 1920, inclusive. The study grew out of the Board's own desire for aid and criticism.

A total of 445,620 marks is embraced in the study. These marks represent four school subjects (English Readings, Elementary French, Elementary Algebra, and Plane Geometry), and cover a period of 19 years. As a basis for comparison, where necessary, the marks were corrected to conform to the following groupings: 90-100, 75-89, 60-74, 50-59, and 0-49. A separate distribution was then made for each subject each year, the work resulting in 76 distributions. Of these, 71 distributions proved to be bimodal with one mode always occurring in the lowest mark-group, and 70 distributions were skewed negatively or toward the low end of the marking scale. When all the marks were thrown into one "grand total" distribution the same bimodality and negative skewing appeared, and only "slight improvement" was noted when the writer prepared a special series of 20 distributions of the marks of "only those candidates who were recommended for examinations on the ground of full and satisfactory preparation."

Comparing these distributions with the normal curve the writer is "amazed at the remarkable extent of divergence." He then outlines the reading methods pursued by the examiners and offers criticisms which need not concern this review. Finally, "the suggestion is made that some approximation to the normal curve offers the best basis for solving present irregularities."

The report might have been made clearer and stronger by including a "grand total" distribution of individual marks from 0 to 100. As it stands, the reviewer is left in doubt whether the divergence from the normal curve, in the actual assignment of marks by the Board, is really as great as is repre-

sented. Examination of the writer's groupings shows that they do not embrace equal ranges of the marks assigned. Thus the first or highest group includes 11 points on the 0-to-100 scale, the second group includes 15 points, the third 15 points, the fourth 10 points, and the last one 50 points. His deductions appear to be made by comparing the percent of marks found in the first group (11-point range) with the percent found in the fifth group (50-point range), and the percent found in the second group (15-point range) with that found in the fourth group (10-point range). The "second mode" of the writer, upon which chiefly seems to rest the finding of "negative skewness," is found in the group which embraces 50 points. Some of the subject distributions show a larger percent of marks in the fourth than in the second group, but comparison of these groups in the "grand total" distribution (p. 8) shows that the latter includes about 33 percent more marks. It should also be noted that the middle group has its lower limit ten points above the mid-point of the scale.

These facts suggest that the varied ranges of the different groups may be the main factor in determining the "grand total" distribution as found, and in its seemingly marked divergence from the normal curve. The same factor appears to enter into many of the subject distributions. Such a table of "raw data" as has been mentioned might have disproved these inferences, but from the report as it stands the Board might well infer that the chief need is to push down the passing-point in its scale to some mark nearer zero.

B. F. Pittenger

University of Texas

News Items and Communications

This department will contain news items regarding research workers and their activities. It will also serve as a clearing house for more formal communications on similar topics, preferably of not more than five hundred words. These communications will be printed over the signatures of the authors. Address all correspondence concerning this department to Doctor E. J. Ashbaugh, Ohio State University, Columbus, Ohio.

The Third Conference of the Bureau of Educational Reference and Research of the University of Michigan was held Tuesday, March 28th. The program consisted of eleven ten-minute reports from workers in the field. The papers were largely grouped about the two subjects of intelligence testing and silent reading, though other phases of school work are not entirely neglected.

Superintendent Bonner Frizzell, Palestine (Texas) reports that for the past three years a local newspaper has given a full page to high-school news. This page is edited by a staff of students under the supervision of the English Department. He also reports that a special study room teacher is employed who has charge of attendance records, monthly reports, and supervision of the study hall, but does no teaching. Such an arrangement is said to be highly satisfactory.

We have received from the Superintendent's office of the Hackensack (New Jersey) Public Schools a Teacher's Manual, a Janitor's Manual, and a Bulletin on Seat Work Exercises for grades one to three. The Teachers' Manual contains the general information with which each teacher in the school system should be familiar. It is in mimeographed form but fully indexed. This fact permits easy revision and adds to the accessibility of the material. Such a manual placed in the hands of a new teacher should assist greatly in helping her to adjust herself to the workings of the system. The Bulletin on Seat Work which was prepared by the teachers should make for uniformity of such work without in the least interfering with initiative in the right direction.

The Normal College Index from the State Normal College, Dillon, Montana for March 10th is called the "School Discipline Number." The main articles in the number deal with the problems of discipline. "How Disciplinary Problems are Solved by Your Co-Educators," "The Relation of Discipline in the School to Community Life," and "Discipline in Theory and Practice" are among the subjects. It is evident that the Normal School is endeavoring through its paper to carry out to the teachers in the State some of the lessons which the instructors are giving the resident students.

Superintendent J. H. Bentley, Duluth (Minn.) Public Schools reports that during the month of February, 1922, which was not a very good month, there was deposited by the Duluth School children $6,733.53 with a total of 18,452 deposits. On January first the total amount on deposit was $146,509.09 with a total of 14,944 accounts. Here is another illustration of the fact that school savings banks enable a considerable amount of the children's money to survive the onslaught of the Holiday Season.

The Los Angeles (California) High School Research Bulletin from which extracts have been made in previous numbers of the Journal continues to carry a large amount of very important material. The issue of March 13th consists of four parts and deals with special articles on subjects of the High School Curriculum by Dr. Franklin Bobbitt. The articles are written as a basis for further work by the Course of Study Committees, and deal with Physical Development and Maintenance, English Experience, Practical Arts for Boys, and Mathematics. In each, specific objectives are set up and principles involved are given consideration. Those who are vitally interested in these problems of the secondary schools and who are willing to use such material should endeavor to secure these bulletins from Los Angeles. We are quite sure that both Dr. Bobbitt and the school people of Los Angeles with whom he is working at present will be glad to have the bulletin used in other school systems.

The Educational Bulletin of the State Department of New Jersey, January, 1922, contains a splendid article on Community Civics and Science by Miss Amelia O. Kays, teacher in the Newton High School.

The major divisions of the outline of materials are (1) The Smaller Community, as the home or the school; (2) The Larger Community—town, city, county, state, federal government; (3) Government as a means to secure cooperation.

Heat, light, water supply, disposal of waste, home conveniences, and recreation are treated under the first head; Public Welfare, transportation and communication, chemistry, come under the second; and the Needs of Government and the industrial and economic problems of government are given consideration in section three.

The outline would be of interest to teachers who have anything to do with the teaching of community civics or science in junior or senior high schools.

The Bureau of Educational Service, Extension Division, State University of Iowa, has just issued its annual report on "Status of Teachers Salaries in Iowa" for the current year. The report is the work of Dr. H. A. Greene, Director of the Bureau, and his Assistant Mr. Don C. Rogers. Data were secured in regard to grade teachers from 812 systems in the state involving over 6,500 positions, and in regard to superintendents, high-school principals and high-school teachers from 720 school systems involving over 4,500 positions. The status of salaries is shown in terms of medians and quartiles for the various grades and positions in different-sized cities. Cities were classified into eight groups on the basis of size. Medians show that in all classes of

cities with the exception of towns of less than 500 population, the salary of grade teachers was greater in 1921-22 than in previous years.

The median salaries of superintendents showed increases for men except in cities of 5,000 to 10,000 population. High-school principals increased their salaries with but two exceptions—namely, in places of less than 1,000 population and in the case of men in places of 2,000 to 5,000 population. Women high-school teachers showed increases in cities of every size. Men high-school teachers, however, increased their salaries only in cities of 2,000 to 5,000 and in cities of over 10,000.

The Lewiston (Idaho) State Normal School Bulletin for January, 1922, is devoted to the Testing and Teaching of Silent Reading, by C. L. Harlan, Head of the Department of Tests and Measurements and his Assistant, I. N. Madsen.

The Bulletin reports results from the Monroe Silent Reading Test in the schools of Idaho during the school year of 1920-21. Over 7,000 children were involved. Distributions and central tendencies are presented in clear detail, and graphs are given to show the relative standing of class A cities, small cities, and rural schools. Three of the conclusions are as follows: On the basis of the central tendencies of all children in each type of schools, the ranks on both speed and comprehension are the same. Schools in Class A cities score best, small city schools next, and rural schools poorest. Intelligence is shown to be a potent factor in the achievement in silent reading. Pupils in any given grade who have spent the fewest years in school have the best scores in both speed and comprehension.

One section of the report is given to suggestions for improvement in silent reading attainments. These suggestions cover recommendations as to how to increase speed or to increase comprehension. The following suggestions are given for meeting the individual needs of pupils: (1) grouping of pupils, (2) personal attention to individuals, (3) gradation of materials used in practice, (4) variation in the amount of reading practice, (5) adaptation to and development of motives, incentives, and interests of individuals, and (6) specific phases of instruction.

Some of our readers may not be acquainted with the Research University of Washington, D. C., which was founded and incorporated under the laws of the District of Columbia for the purpose of providing federal employees and other men and women with an opportunity of gaining such college training as is necessary for efficiency and happiness. Most of the work is taught after five o'clock in the evening when day workers may take advantage of it, although day classes are also provided in a number of subjects. About 200 correspondence courses are offered for those out of the District. The institution has grown from one instructor, a handful of students, and one room in the fall of 1919 to a faculty of 75, a student body numbering nearly 1,000, and two large buildings. The president of the institution is Dr. Louis W. Rapeer.

The University is operated on a cooperative basis. The tuition is kept as low as is consistent with the policy of avoiding begging either from indi-

viduals or from the government, and the institution shares the income with the teachers on a 50-50 plan.

The institution is doing college and university work, requiring the standard 15 units for entrance and 120 hours for graduation, and conferring regular degrees upon graduates. It is the only institution of higher learning in the Capital which grants the degree of Bachelor of Secretarial Science. For the degree of Bachelor of Commercial Science, three years of successful business experience are required in addition to the regular college work.

Cuyahoga County is one of the few counties in the country which has a full-time Assistant Superintendent in charge of Educational Research work. Bulletin No. I from the County Bureau contains 42 pages of mimeographed material dealing with tests given to about 6,600 children in 31 different districts of the county. The tests used were the Illinois General Intelligence Scale, Burgess' Silent Reading Scale, Monroe's General Survey Scale in Arithmetic, Charter's Diagnostic Language Tests, and for Spelling 25 words for each grade from the Buckingham Extension of the Ayres' Spelling Scale. The tests are being given twice the first year, and the following year the teachers are expected to give these or similar tests three times.

The Bulletin contains an introduction dealing with the value of measurement. The report on the results of the tests is written in an exceedingly clear and simple manner. However, the report does not stop at that. It gives definite suggestions for improvement. The various districts are finally ranked on the basis of an educational index number. This index number was found by adding one-third of the achievement quotient in reading plus one-fourth of each of the achievement quotients in arithmetic and language, plus one-sixth of the achievement quotient in spelling. This relative weighting of the different subjects was on the basis of the judgment of the local superintendents who were asked to rate the importance of these subjects.

Superintendent Yawberg reports that the school boards and the public generally in the various districts have shown great interest in the measuring and ranking and that one member of the board of the district ranking lowest said: "We may be lowest in this first measurement, but if we are lowest in the next, we will find out the reason why." That spirit is prophetical of activity in the right direction.

Superintendent Herman Dressel, Kearny, New Jersey, reports the following interesting experiment which is now being conducted in his schools. We shall certainly look forward with interest to the later report on the outcome of the experiment.

"Pupils know too little of the teacher's plans. We have trained them into a ready acceptance of whatever we may dole out to them and they have, in a submissive spirit accepted their lot and their daily portion of work to be done. If the course of study were in the hands of the pupil as well as in the mind of the teacher, he would have a little clearer idea of where he is going, and whether he is on his way. With that idea in mind, several of the Principals began to plan for a series of lessons. These were intended to give

the pupil certain information and certain directions, and to place before him definite tasks to be done. In the subject of arithmetic these tasks would be graded examples. On each page he finds 'The Way To Do It.' This consists of a typical example worked out and explained. The pupil then tries his hand at working the examples given. On the sheet are directions to look for others in textbooks. The sheets we have called Job Sheets, because we are bunching them into pads of eight and ten; and when the pupil finishes a set he gives this to his teacher who places it with his other work as a finished job. When he has completed his entire set of jobs, he will be given others. In this way he is working ahead of his class, if he is a bright pupil, and trying to catch up, if he is backward. We are at present working out the job sheets in English and arithmetic. We feel confident that the job sheets will prove a time saver and will also demonstrate to teachers that pupils can do a great deal more work than we have imagined them capable of doing. For instance, a principal and I both decided that a certain number of jobs which we had ready would require two weeks for the pupils to finish. Without any discussion as to time and with the simple suggestion that this work might be done whenever they had any extra time after preparing their regular lessons, the pupils who accepted the sheets finished the entire set in two days! Since then my office force has tried in vain to keep up with the demand that the pupils are making for the job sheets.''

National Council for Social Studies

A National Council for the Social Studies completed its organization in Chicago on February 25th. Its purpose is to lay the foundations for training democratic citizens; and its sponsors believe that such training can result only from a carefully developed and adequately supported system of teaching in the elementary and secondary schools. Its plan looks to promoting cooperation among those who are responsible for such training, including at least the university departments which contribute knowledge of facts and principles to civic education; and the leading groups of educational leaders, such as principals, superintendents, and professors of education, who develop the methods of handling these facts.

An advisory board was set up composed of representatives of (1) the five associations of scholars most nearly related to the purpose of the National Council,—historians, economists, political scientists, sociologists, and geographers; (2) the national organizations of educational investigators and administrators—elementary and high school principals, teachers of education, normal school principals, and superintendents; and (3) regionary associations of teachers of history and civics. The function of this advisory board is to bring into the National Council the points of view of the organizations represented by its members and to insure a development of the social studies which will be in harmony with the best educational thought as well as based on the best present practice.

The following officers were elected for the year 1922-1923: L. C. Marshall, Professor of Economics in the University of Chicago, President; Henry Johnson, Professor of History in Teachers College, Vice-President; Edgar Dawson,

Professor of Government in Hunter College, Secretary-Treasurer; E. U. Rugg, Lincoln School, New York, Assistant Secretary. An executive committee, charged with the general direction of the policies of the association will consist of the officers and the following elected members: C. A. Coulomb, District Superintendent, Philadelphia; W. H. Hathaway, Riverside High School, Milwaukee; Bessie L. Pierce, Iowa University High School.

The first task the National Council is undertaking is the preparation of a Finding List of those experiments or undertakings in the teaching of the social studies which now give promise of being useful. This list will contain such exposition of the character and aims of these experiments as to make it possible for those working along parallel lines to discover each other and to cooperate more fully than would otherwise be probable. This expository material will have another purpose,—that of indicating outstanding differences of opinion and program in order that these differences may be systematically stated for purposes of analysis and discussion.

To aid in the discovery and assessment of these experiments, the National Council has in preparation a list of *Key Men and Women* who will be appointed in the various states to represent the National Council in its efforts to collect useful information and then to give currency to it. While this organization seems to represent all the elements out of which the best development of the social studies must proceed, the most useful work will be done only with the cooperation of teachers and investigators in all parts of the country to the end that lost motion and useless repetition may be eliminated and that mutually strengthening experiments may be pressed forward.

Persons who are interested in the wholesome development of the social studies, whether teachers or others, and if teachers, whether teachers of the social subjects or of some other subject, are urged to communicate at the earliest convenient moment with the secretary of the National Council, Edgar Dawson, 671 Park Avenue, New York City.

Ohio State Educational Conference

The second annual meeting of the Ohio State Educational Conference was held at Ohio State University March 23-25, 1922. The program consisted of three general sessions and twenty sectional sessions. The general theme of the Conference was Supervision, with special emphasis upon the place of the Building Principal in the supervisory program. The keynote of the Conference itself was expressed by Dean George F. Arps of the College of Education, who as presiding officer at the first session, expressed the hope that the Conference would be a *working conference*. This hope was transformed into a reality, for every paper presented workable material and the response from the floor whenever discussion was permissible showed that the audience was quite willing to contribute to the success of the program.

Among the speakers who addressed the general sessions were: President W. O. Thompson, Dean George F. Arps, Dr. B. R. Buckingham, Dr. Boyd H. Bode, all of Ohio State University; Superintendent R. G. Jones of Cleveland and Superintendent R. J. Condon of Cincinnati; A. C. Perry, Jr., District Super-

intendent, and Angelo Patri, Principal P. S. 45, New York City; and Dr. W. W. Charters, Carnegie Institute of Technology.

The sectional meetings which occurred on Friday included the following: Administrative and Supervisory Research, High School Principals, Elementary Principals, Teacher-Training, American History, Art, Biological Science, Non-Biological Science, English, Latin, Mathematics, Modern Languages, Home Economics, Industrial Arts, Music, Vocational Agriculture, Physical Education, Clinical Psychology, Educational and Psychological Tests, and a demonstration of Educational Tests.

It had been hoped by those in charge that the program would appeal to all classes and grades of educational workers in the State. The attendance of over 2,500 persons representing practically every phase of educational work would seem to justify the belief that this purpose was realized. A number of cities dismissed school in order that all teachers might be present for the Friday and Saturday sessions. Others made special provisions for the release of the administrative and supervisory forces.

The exhibits held during the Conference attracted a great deal of attention. Professor A. E. Waller of the Department of Botany and Professor D. M. DeLong of the Department of Zoology and Entomology prepared an exhibit of biological material which indicated ways of improving teaching and increasing interest in the biological sciences. The Department of Physical Education of the University displayed what was, as far as can be learned, the first physical education exhibit to include all phases of such work from the elementary school through the university. The Department of Psychology had an extensive exhibit of educational and mental tests. An exhibit of some two thousand textbooks and educational pamphlets was prepared by the Bureau of Educational Research. The large numbers attending these exhibits and the great interest evinced by all visitors was satisfactory evidence that a distinct service was rendered in this manner.

The Validity of Timed-Sentence and Column Tests in Spelling

A few years ago a number of workers in educational research, notably Courtis, Fordyce, and Monroe, called attention to the fact that "the percent of correct spellings is higher when the words are dictated separately than when the words are dictated in timed sentences and the pupils forced to write at their normal rate."[1] The subsequent production of "Timed-sentence Spelling Tests" has greatly affected the practice of surveyors, superintendents, and teachers desiring to measure the spelling ability of groups of pupils. The prevalent opinion seems to be that "by using the words embedded in sentences and dictated rapidly enough to force the child to write at his normal rate, we measure the spelling ability which functions in one's every-day writing."[2] The Monroe tests and similar tests constructed by teachers and superintendents have been very extensively used. I have before me five mimeographed bulletins reporting the results of this year's (1920-21) testing in as many Ohio cities. All employed timed-sentence tests for measurements in spelling.

[1] Monroe, Walter S., *Measuring the Results of Teaching*. Boston: Houghton Mifflin Company, 1918. p. 182.

[2] Ibid, p. 183.

The complaint, however, is frequently heard that timed-sentence spelling tests are difficult to score and the resulting scores less accurate than those obtained from tests of the conventional column type. Investigation has shown that nearly twice as much time is required to score timed-sentence test papers as is necessary for a column test containing the same number of test words. Perhaps, therefore, it is in order to inquire as to the relative validity of these two kinds of tests. Granting that pupils do not do as well on timed-sentence as on column tests, may it not be true that there is a high degree of consistency between the results of the two? If, in short, there is a high correlation between the results of two tests, one timed-sentence and the other column, it is fair to question the wisdom of expending the time and energy involved in preparing and administering timed-sentence tests and scoring the papers.

However high the correlation it cannot be adequately interpreted without some basis for comparison. Accordingly one timed-sentence and two column tests were prepared in order that the correlation between the scores of the two column tests might be compared with that between either of the column tests and the timed-sentence test.

Letters were addressed to five Ohio city superintendents[1] who were known to be interested in the problems of educational research. All very cordially agreed to give the tests to sixth-grade pupils and to supervise the scoring of papers and recording of scores. The tests are given below.

SPELLING TESTS FOR SIXTH-GRADE PUPILS

Instructions

The three tests should be given in succession to a group of pupils. All sixth-grade classes in the system should be tested on the same day, if possible, or, if this is impossible, on two successive days. The days are Thursday, December 2, and Friday, December 3, 1920.

Test A

Say to the pupils, "I have here a list of words which I want you to spell. You will need a large sheet of paper and a pencil or pen. Better have an extra pencil within reach, so that if you break one you will not be delayed." Pause until all are ready. Then say, "Write your name at the top of the sheet. Number your words down the left hand side 1 to 25. (Pause) I shall pronounce each word twice. I shall not use the words in sentences nor tell you what they mean, so do not ask questions. Take time to dot your i's and cross your t's. All ready!"

The list follows:

1. lose	6. decide	11. invitation	16. opinion	21. total
2. machine	7. wreck	12. accident	17. celebration	22. supply
3. improvement	8. busy	13. success	18. minute	23. accept
4. assist	9. particular	14. different	19. forenoon	24. official
5. organize	10. therefore	15. believe	20. avenue	25. complete

[1] Grateful acknowledgment is made to Supt. J. V. Nelson, Bellaire; Supt. J. R. Patterson, Bucyrus; Supt. F. J. Prout, Chillicothe; Supt. E. L. Daley, Circleville; Supt. R. C. Maston, Martins Ferry; and Mrs. Margaret Brainerd, Director of Educational Research, Martins Ferry.

TEST B

"Now turn the sheet over and use the other side. 'I have some sentences which I want you to write as I dictate them. I am going to dictate them rather rapidly, possibly more rapidly than some of you can write. If you have not finished writing one sentence when I begin to dictate another, I want you to leave it and begin on the new sentence. If there are any words you cannot spell, you may omit them. Take time to dot your i's and cross your t's. If you have any question about what you are to do, ask it now, because you cannot ask questions after I begin to dictate.'"[4]

The sentences follow:

60 *Christmas* came on *Wednesday* last year.
32 *Really* you *ought* to *refer* the *argument* to your teacher.
17 His *condition* is *serious* today.
43 *Education* is for the *common* people.
12 They will *arrive* by *automobile tomorrow* night.
51 *Already* the *theater* is showing *motion* pictures.
32 He will *entertain* a *visitor* next week.
3 I am *often* in *doubt* about my lesson.
31 The *recent conference* was a *pleasant* one.
5 The *general purpose* is good.

When the second hand reaches 28 stop the writing.

TEST C

"Now turn the sheet over again. I want you to write another list of words beside the list which you have already written. I shall pronounce each word twice, as I did before. Remember to take time to dot your i's and cross your t's and to ask no questions."

The list follows:

1. combination	6. circular	11. estimate	16. system	21. government
2. neighbor	7. article	12. impossible	17. associate	22. witness
3. difference	8. increase	13. treasure	18. political	23. national
4. examination	9. neither	14. director	19. attention	24. business
5. volume	10. marriage	15. local	20. service	25. convention

The class record sheet provided blanks for the names of the pupils and the three scores in adjacent columns. It was therefore a simple matter to correlate the scores. The Ayres short-cut for the Pearson formula[5] was employed for this purpose. The statistical work is considerably lightened by this procedure provided that one possesses such mechanical aids as a listing adding machine and calculating tables. All work was checked to insure a high degree of accuracy. On the average, one error was found in each 2500 items.

[4] Monroe, op. cit. p. 185.

[5] Ayres, Leonard P., "A Shorter Method for Computing the Coefficient of Correlation," *Journal of Educational Research*, 1: 216, 221.

The results are presented in the accompanying table. For the 275 sixth-grade pupils at Bellaire the value of r for the scores of Tests A and B is 0.82; for Tests A and C it is 0.87; and for Tests B and C, 0.87. That is, the scores on the timed-sentence test, Test B, correlate as highly with the scores on one

City	Number of Cases	Value of r for		
		A & B	A & C	B & C
Bellaire	275	0.82	0.87	0.87
Bucyrus	147	0.86	0.88	0.87
Chillicothe	218	0.89	0.89	0.87
Circleville	115	0.82	0.93	0.85
Martins Ferry	252	0.81	0.87	0.87
Total	1007	0.83	0.89	0.87

of the column tests, Test C, and almost as highly with the scores on the other column test, Test A, as do the scores of one column test with the scores of the other. Similar statements may be made for the remaining four cities, with the exception of Circleville, the smallest system included in the study.

It may be contended that we are loading the dice against the timed-sentence test by having two correlations of the scores of column tests with those of the timed-sentence test and only one between the scores of the two column tests. Perhaps a fairer statement of the case may be made by averaging the first and third coefficients for each city and comparing this average with the second coefficient. This results in each case, in a slight margin in favor of the timed-sentence test, a margin of four points for the entire 1007 cases in the five cities.

In the judgment of the writer this slight increase does not justify the expenditure of time and energy required for timed-sentence tests, so far as their use for survey or measurement purposes is concerned.

R. L. Morton

Ohio University, Athens, Ohio

Mental Ability Levels of Classes Versus I. Q. in Measuring Educational Output

The practice of evaluating the relative intelligence of classes in order to measure the efficiency of teachers or of methods of instruction is becoming more prominent in our schools. This is a very hopeful sign for those who labor to put educational procedure on a more scientific basis. Attempts to arrive at useful conclusions in this way may, however, be very easily frustrated by incorrectly applied principles. One of these erroneous methods consists in expressing the intelligence of a class in terms of average or median intelligence quotient instead of mental age. The former is invalid as a measure of mental ability in supervision save where the average chronological ages of classes are the same. Mental age, on the other hand, indicates the level of mental growth that a class has reached and this is the important thing from the standpoint of ability to master subject matter in school studies. The average

I. Q. indicates the brightness of a class, but tells nothing of how high or how low it stands in respect to mental development. A first-grade class and an eighth-grade class may both have an average I. Q. of 110 which places them on a par so far as brightness is concerned. The difference in intelligence, however, would be about seven years of mental age. This difference is the significant one for purposes of supervision.

In Saint Paul we have just concluded the examination of thirty-two VIII-A classes. In each case we have calculated the median mental age and the median I. Q. If one arranges these classes according to mental age, the order will be very different from that of the classes ranked according to I. Q. For example, the class occupying the second position from the top drops to the seventeenth place. Other classes rise or fall a number of positions. This is, of course, due to the varying chronological ages of the pupils who make up the several classes. To rate according to the I. Q. is to assume that the average ages of classes are the same, but this assumption is not true and the conclusions drawn from it are not valid. The table given below shows the displacements in rank which take place when classes are rated according to median I. Q. instead of median mental age. It should be remembered that the maximum possible displacement in rank among 32 classes would be 31 positions.

Displacements in rank	Number of Classes
0	1
.5	1
1	1
1.5	4
2	5
2.5	3
3	1
3.5	1
4	1
4.5	2
5	2
5.5	1
6.5	2
7	1
8	2
10	1
11	1
14.5	1
Average	4.2

The above table shows that only one class out of 32 retained the same rank. Several classes suffer severe changes, while the average of the displacements is 4.2 positions. This shows clearly that these two methods of evaluation of mental ability will yield widely different results and that as a consequence workers in mental measurements should be careful which one they use if they would secure reliable data.

O. J. JOHNSON
Ass't Director Research
St. Paul Public Schools

Grading by Means of Standardized Tests

Mr. L. C. Wilkerson, Principal of the Opportunity Farm School, Cincinnati, Ohio, has sent us a brief statement of his efforts at the proper grading of 75 boys, all delinquents. No school records accompany the boys and their attendance is limited to one-half of each school day. The psychology tests given by the Vocational Bureau showed only four boys with an I. Q. of one hundred or better. Standard educational tests were given in reading, vocabulary, spelling, composition, arithmetic, geography, and history. The records of each individual were graphed in terms of standard grade achievements, as illustrated in the accompanying figure.

INDIVIDUAL GRAPH RECORD

Name, J.W. Born, Sept. 5, 1906 Grade V

Mental Retardation or Acceleration, 1 yr. 9 mo. I. Q., 88

Test	Grade					
	3	4	5	6	7	8
Monroe Reading, Rate					*	
Monroe Reading, Composition			*			
Courtis Reading, Rate						*
Courtis Reading, Questions	*					
Courtis Reading, Index			*			
Gray, Oral Reading			*			
Thorndike, Vocabulary						*
Spelling, List		*				
Spelling, Timed Sentence		*				
Geography					*	
Composition, Story		*				
Composition, Form		*				
Arithmetic		*				

As a result of these tests, reclassification was made. Out of each ten, one could be advanced one grade, four were properly placed, four belonged one grade lower, and one two grades lower.

As the result of this reclassification, a new school spirit was created. Each boy's graph showed where he needed help and each teacher taught in terms of the need. The boys were anxious to take tests and record the progress they were making in their various studies. After a year's trial it was found that "specific treatments for specific educational ailments produced some striking results." Tests in silent reading at the end of the year showed each grade, with few exceptions, at or above standard. All boys at the end of the year were promoted except two who had received promotion at mid-year. Considering the fact that school work occupied but one-half of each day, the results were most gratifying.

Educational Research Association

(FRANK N. FREEMAN, *Secretary and Editor*)

It has been the custom to print in this section personal notices concerning members of the Association, and brief accounts of the problems, on which they are working or of the results of investigations which are not extensive enough to justify writing articles upon them. On the account of the election of a new secretary at the last meeting, the members of the Association have not yet acquired the habit of sending him material. This is a renewal of the standing invitation to send appropriate notices, or literature regarding important meetings, conferences, etc., or concerning scientific investigations which are in progress or completed. Let each member send the kind of information about himself that he would like to have about his fellow members. Address the secretary at the University of Chicago.

The secretary proposes to throw the section open to another type of communication. It seems to him that it would be profitable to make it a forum for brief discussions of problems connected with scientific investigations. The problems might be concerned with the technique of investigation or with the interpretation of its results. They might be general in their reference or they might refer to some particular published study. They should in all cases be brief, pointed and specific. Administrative problems arising out of the conduct of research or of the scientific survey of school systems might be treated. Requests for information on proposed investigations would also be appropriate. The suggestion is that the Journal be used as the medium for a continuous round-table discussion of our professional problems.

Leon O. Smith, Assistant Superintendent of Schools of Omaha, will give a course in Public School Publicity and one on The Elementary School Curriculum at the University of Montana during the second term of the coming summer session. Mr. Smith has recently been elected President of the Nebraska Conference of Social Workers.

A conference of the Bureau of Educational Reference and Research of the University of Michigan was recently held under the direction of Dr. Clifford Woody. The program was contributed to by administrators and teachers from various parts of the state and consisted chiefly of reports of the use of mental or educational tests. Some of the papers will be published in this journal. The meeting was well attended and much interest was manifested in the display

of educational and intelligence tests and in the exhibit of numerous graphic and tabular methods of presenting the results of tests.

A report has been received of the results of a series of United States history tests in Boston. Both the tests and the investigation were organized by Miss Olivia C. Penell, Research Assistant in the Department of Educational Investigation and Measurement. A test is designed for each of Grades VI, VII, and VIII. Some of the tests are common to three and some others to two of the grades. The questions are largely of the type which test information.

TABLE I—COMPARATIVE RESULTS ON QUESTIONS COMMON TO GRADES VIII, VII, VI

Grade		VIII		VII		VI	
Question	Possible Credit	Median Credit	Per-cent of Accu-racy	Median Credit	Per-cent of Accu-racy	Median Credit	Per-cent of Accu-racy
I. Period of Discovery and Exploration	10	9.1	82.0	9.04	80.3	8.95	79.8
II. Period of Settlement: 1. Associate colony with Nation and date	10	5.6	51.3	5.6	51.3	6.9	59.8
2. Associate man with colony	5	3.2	55.4	3.1	52.2	4.2	67.3
3. Causes of settlement	2	1.3	41.1	1.3	38.4	1.5	50.9
Patriotic Dates	10	6.9	59.7	6.7	58.8	6.1	54.3

TABLE II—COMPARATIVE RESULTS ON QUESTIONS COMMON TO GRADES VIII AND VII

Grade		VIII		VII	
Questions	Possible Credits	Median Credits	Percent of Accuracy	Median Credits	Percent of Accuracy
III. Pre-Revolutionary and Revolutionary Period	10	5.7	50.8	6.1	55.98
Inventors and Inventions	10	6.02	55.2	3.95	37.9
Late Events	3	1.96	48.5	1.1	30.8

Some require the underlining of one out of several items, some for supplying words, some for the completion of sentences. The information called for consists of dates, names of persons and places connected with important events, facts concerning political, social, and industrial movements, etc. The tests were given to over 900 pupils of each of the three grades. Tables I and II show the comparative scores for the different grades.

Among the conclusions were the following. Many of the historical facts which are taught in the sixth grade are remembered during the two remaining grades. (Nothing is said in this connection concerning the possible effect of elimination.) Fact questions obtained a higher percent of accuracy than judgment questions.

SUCCESS IN SCHOOL WORK

Should Exceed
That Due to Native Ability

Walter S. Monroe is reported as saying that the dull pupil does better class work and makes a better school record than the bright one. Assistant Superintendent E. K. Sexton confirms this.

The Illinois Examination measures the success of each pupil by his own ability to do school work. It shows what the school accomplishes by the instruction given in the fundamentals of arithmetic and of silent reading.

THE RESULTS ARE SURPRISING

About 500,000 pupils have taken *The Illinois Examination*. Almost every state in the Union has used it.

The returns show that everywhere the school does more for the pupils of average or less than average ability than for gifted children.

WHAT TO DO AFTER USING THE ILLINOIS EXAMINATION

After you have used *The Illinois Examination* in your school, the scores of the pupils furnish a basis on which to plan improvement in the curriculum and in individual instruction.

There is now little question but that, under proper incentives, all children can increase their capacities to do the work of the elementary school. The *Illinois Examination* measures such increase in capacity. If the score of one boy in a class shows an achievement quotient of 123, there is every reason to expect that eventually every member of the class can be brought to attain the same high rank in mastery of school work.

In ordering the *Illinois Examination* specify Examination I for grades 3, 4 and 5, and Examination II for grades 6, 7 and 8.

Both Examinations I and II are furnished in two forms so that, if desired, the same pupils may be tested two times during the school year.

The *Illinois Examination* is used successfully as a final examination at the end of the spring semester or term.

General Intelligence and Achievement Combined

1. Illinois Examination. This includes a General Intelligence Test and Tests in Operations of Arithmetic, and in Silent Reading. Standardized from 55,000 scores.
   Form 1. Sample Set, including Teachers Handbook, 25 cents .....
   Illinois Examination I, for grades 3, 4 and 5, $4.00 per 100 (16 pages) .....
   Illinois Examination II, for grades 6, 7 and 8, $4.00 per 100 (16 pages) .....
2. Form 2. Sample Set, including Teachers Handbook, 25 cents .....
   Illinois Examination I, for grades 3, 4 and 5, $4.00 per 100 (16 pages) .....
   Illinois Examination II, for grades 6, 7 and 8, $4.00 per 100 (16 pages) .....

Teachers Handbook separate by mail, 15 cents each .....

Each Test of the Illinois Examination is also printed separately.

Orders Filled Promptly

PUBLIC SCHOOL PUBLISHING CO.
BLOOMINGTON, ILLINOIS

THE USE OF STANDARDIZED ARITHMETIC TESTS

Thousands of schools are determining the value and extent of arithmetic instruction by using the standardized arithmetic tests listed below. Armed with the results obtained from the use of these standardized examinations, the superintendent and principal are in a better position to judge both the ability of the teacher and the value of the text used. The tests by Monroe and Buckingham are in constant demand all over the United States. The Monroe General Survey Arithmetic Test is particularly valuable because it is furnished in three forms so that the same pupils can be tested three times during the school year if necessary. The Cleveland Survey Arithmetic Test is also a diagnostic test.

Buckingham Scale for Problems in Arithmetic
- Form 1
- Sample Set, 8 cents
- Division I, for grades 3 and 4, 80 cents per 100
- Division II, for grades 5 and 6, 80 cents per 100
- Division III, for grades 7 and 8, 80 cents per 100
- Form 2
- Sample Set, 8 cents
- Division I, for grades 3 and 4, 80 cents per 100
- Division II, for grades 5 and 6, 80 cents per 100
- Division III, for grades 7 and 8, 80 cents per 100

Monroe General Survey Arithmetic, See Illinois Examination, (page I of this magazine.)
- Form I. Scale I Monroe's General Survey Arithmetic, for grades 3, 4 and 5, $1.00 per 100
- Scale II Monroe's General Survey Arithmetic, for grades 6, 7 and 8, $1.00 per 100
- Form 2. Scale I Monroe's General Survey Arithmetic, for grades 3, 4 and 5, $1.00 per 100
- Scale II Monroe's General Survey Arithmetic, for grades 6, 7 and 8, $1.00 per 100
- Form 3. Scale I Monroe's General Survey Arithmetic, for grades 3, 4 and 5, $1.00 per 100
- Scale II Monroe's General Survey Arithmetic, for grades 6, 7 and 8, $1.00 per 100

Cleveland Survey Arithmetic Tests, for grades 3 to 8, $1.90 per 100
- Sample Set, 10 cents

Monroe Diagnostic Tests in Arithmetic
- Sample Set, 10 cents
- Part I, Integers, for grades 4 to 8, 85 cents per 100
- Part II, Integers, for grades 5 to 8, 85 cents per 100
- Part III, Common Fractions, for grades, 6 to 8, 85 cents per 100
- Part IV, Decimal Fractions, for grades 6 to 8, 85 cents per 100

Monroe Standardized Reasoning Tests in Arithmetic
- Form 1
- Sample Set, 8 cents
- Test I, for grades 4 and 5, 80 cents per 100
- Test II, for grades 6 and 7, 80 cents per 100
- Test III, for grade 8, 80 cents per 100
- Form 2
- Sample Set, 8 cents
- Test I, for grades 4 and 5, 80 cents per 100
- Test II, for grades 6 and 7, 80 cents per 100
- Test III, for grade 8, 80 cents per 100

Orders Filled Promptly

PUBLIC-SCHOOL PUBLISHING CO.
BLOOMINGTON, ILL.

Otis Self-Administering Tests of Mental Ability

By Arthur S. Otis

These tests are a distinct advance in mental ability testing. They embody four new features: 1. "Self-administration," so that the examiner has but to give a few initial directions; 2. Ease of scoring, which reduces the time for scoring a paper to about 45 seconds; 3. Flexible time limit, so that the tests can be given with a time limit of either twenty or thirty minutes; 4. Variety of test material, consisting of some 24 types of questions that insure a more comprehensive measurement.

The *Higher Examination* is designed for grades 9 to 12 and for college freshmen. The *Intermediate Examination* is for grades 4 to 8. Each examination is in two forms—A and B, alike in form and difficulty but different in content.

Higher Examination: Form A. 4 pages. Price per package of 25 examination booklets, 1 Directions and Key, and 1 Record Sheet $1.00 net.

Higher Examination: Form B. 4 pages. Price per package of 25 examination booklets, 1 Directions and Key, and 1 Record Sheet $1.00 net.

Intermediate Examination: Forms A and B. In preparation.

Write for free sample of Higher Examination.

Ready Soon

An Introduction to the Use of Standard Tests

By Sidney L. and Luella C. Pressey

A simple and practical manual making clear the fundamentals of effective testing It has been made as non-technical as possible, and is well adapted to the needs of the busy teacher, supervisor, or superintendent beginning a testing program.

WORLD BOOK COMPANY

Yonkers-on-Hudson, New York 2126 Prairie Avenue, Chicago

READING SCALES APPLIED TO PARTICULAR SUBJECTS

The Van Wagenen Scales in { History / General Science / English Literature }

School authorities find a need for the measurement of the ability to read and comprehend particular school subjects.

The Van Wagenen series of reading scales will meet just this need.

The Van Wagenen scales aid in determining the degree of difficulty of reading material that may safely be placed in the hands of students of various degrees of capacity.

The Van Wagenen scales make possible

(1) relatively accurate comparisons of ability in reading various subjects;

(2) relatively accurate measurements of the effects of increased range of information in a subject upon a student's ability to read intelligently in that subject;

(3) relatively accurate measurements of the effect of different methods of study upon ability to read independently in a certain subject.

The Van Wagenen Reading Scales *are purely reading scales.* No writing other than the making of check marks is necessary.

Added to the ease with which the exercises may be scored is the ease with which the student's score may be calculated in terms of a score value.

There are duplicate scales for each of the subjects so that accurate comparisons of achievement may be made.

Comparisons of achievement in terms of definite degrees of ability may be made, a process that is difficult with the use of ordinary standard tests.

THE VAN WAGENEN READING SCALES

History Scale A (8 pages), or Scale B, price $3.00 per 100.
General Science Scale A (8 pages), or Scale B, price $3.00 per 100.
English Literature Scale A (8 pages), or Scale B, or Scale C, price $3.00 per 100.

THE PUBLIC SCHOOL PUBLISHING CO.
BLOOMINGTON, ILLINOIS

OHIO STATE UNIVERSITY

Announcements

COLLEGE OF EDUCATION

SUMMER QUARTER - JUNE 19 to SEPT. 2, 1922

First Term - June 19 to July 26, 1922
Second Term - July 27 to Sept. 2, 1922

FOUR-QUARTER PLAN

The Ohio State University has adopted the quarter plan of organization. The Summer Quarter in the College of Education is divided into two terms for the convenience of teachers who elect to do a half-quarter work.

The Summer Quarter is prepared to meet the needs of School Superintendents, High School Principals, Normal School Teachers, Supervisors, Junior High Principals, Grade Principals, High School Teachers, Elementary School Teachers, Teachers of Industrial Education, Art Teachers, Music Teachers.

DEGREES

The quarter plan of organization makes it possible to obtain the undergraduate degree (B. S.) and the advanced degrees (A.M. and Ph.D.) by attending a succession of Summer Quarters.

ACADEMIC SUBJECTS

Undergraduate and graduate content courses will be offered to high school and grammar school teachers in practically all of the accepted high school subjects. The University quarter plan of organization makes possible the expansion of the work in academic subjects.

PROFESSIONAL SUBJECTS

Twenty courses will be offered in Principles and Methods of Teaching; fifteen in Psychology; ten in Industrial Education; six in History of Education; fifteen in School Administration; sixteen courses for men and women in Physical Education; four courses in Art; nine courses in Music.

BUREAU OF EDUCATIONAL RESEARCH

The recently established Bureau of Educational Research under the Directorship of Dr. B. R. Buckingham, formerly of the University of Illinois and Editor of the Journal of Educational Research, will be placed at the service of school superintendents, principals, and teachers during the summer quarter.

For information address

The University Examiner

THE OHIO STATE UNIVERSITY
COLUMBUS, OHIO

WITHAM'S TESTS
STANDARD GEOGRAPHY

The only practical diagnostic tests yet devised in Geography.

English Vocabulary—Tests 1 and 2 now available
Silent Reading—Test No. 1. Mechanical Content
—Test No. 2. Literary Content

All of the above tests put up in sets of 50 of one kind complete, with printed directions for giving and scoring tests, standard scores and key.

We also publish

BROOKS GRAPH RECORD CARDS and
MAXSON'S SELF KEYED NUMBER WORK CARDS

Write Dept. A for Complete Sample Sets

J. L. HAMMETT COMPANY
Cambridge, Massachusetts

AMERICAN BOOK COMPANY

EDUCATIONAL PUBLISHERS

NEW YORK, 100 Washington Square
CINCINNATI, 300 Pike Street
CHICAGO, 330 East 22nd Street
BOSTON, 63 Summer Street
ATLANTA, 2 North Forsyth Street

Address all communications to the nearest office of the company. Descriptive circulars, catalogues, and price lists will be mailed on request to any address.

Your correspondence is solicited and will have prompt attention.

A TEST { OF ABILITY TO LEARN and OF ACHIEVEMENT

Education is, as Bagley says, "training for achievement." And Intelligence, according to the definition offered by Buckingham and accepted by many leaders, is "the ability to learn."

Judged by these two standards a test which furnishes you reliable data concerning your pupils' ability to learn and the record of achievement they have made in their school work is the test you wish to use in your school system.

THE ILLINOIS EXAMINATION is such a test. The first seven pages contain the seven sections of the Illinois General Intelligence Test; the remainder of the pamphlet includes the Monroe Standardized Silent Reading Test and the Monroe Survey Scale in Arithmetic.

THE ILLINOIS EXAMINATION was the first test to furnish teachers a means of determining the pupil's "achievement quotient," which is now generally regarded as an essential part of the general information that the teacher should have regarding every pupil under her charge.

THE ILLINOIS EXAMINATION is used with entire satisfaction and success in both city and rural schools. It can be, and is, administered by the teachers themselves without the necessity of calling in an expert.

Though this three-barrelled test is one of the very best to be had, it is also one of the most economical in price. The cost is $4.00 net per hundred copies. We fill orders promptly.

PUBLIC SCHOOL PUBLISHING COMPANY
BLOOMINGTON, ILL.

Three Notable Groups

Science

Black and Conant	Practical Chemistry
Black and Davis	Practical Physics
Good	Laboratory Projects in Physics
Hegner	Practical Zoology
Trafton	Science of Home and Community
Williams	Healthful Living
Winslow	Elementary Industrial Arts

Mathematics

Ford and Ammerman	First Course in Algebra
" " "	Second Course in Algebra
" " "	Plane Geometry
" " "	Solid Geometry
Hayward	Progressive Problems in Bookkeeping and Accountancy
Kiggen	Practical Business Arithmetic

Teacher's Books

Bode	Fundamentals of Education
Book	Intelligence of High School Seniors
Boraas	Teaching to Think
Davis	The Technique of Teaching
Thorndike	Psychology of Arithmetic

Complete Catalogs

THE MACMILLAN COMPANY

New York Atlanta Chicago
Boston Dallas San Francisco